

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	484	100.0	216	2	Q14286	Q14286 homo sapien
2	484	100.0	2351	1	FAS HUMAN	P00451 homo sapien
3	417	86.2	2343	2	O18806	O18806 canis famil
4	412	85.1	2343	2	O62730	O62730 canis famil
5	406	83.9	2319	1	FAS MOUSE	O62194 mus musculu
6	392	81.0	2133	1	FAS PIG	P12263 sus scrofa
7	351	72.5	2258	2	O77N96	O77N96 rattus norv
8	291	60.1	1377	2	O804X3	O804X3 gallus gall
9	240	49.6	1639	2	O804W6	O804W6 figu rubrip
10	208	43.0	427	1	MFGM RAT	P70490 rattus norv
11	206	42.6	116	2	O8C7H9	O8C7H9 mus musculu
12	206	42.6	2183	2	O88783	O88783 mus musculu
13	205	42.4	463	1	MFGM MOUSE	P21956 mus musculu
14	204	42.1	2102	2	O77P6K2	O77P6K2 rattus norv
15	203	41.9	745	2	O804X4	O804X4 gallus gall
16	202	41.7	2258	1	FAS PIG	Q951P1 sus scrofa
17	199	41.1	2224	1	FAS HUMAN	P12259 homo sapien
18	199	41.1	2224	2	O6UFU6	O6UFU6 homo sapien
19	198	40.9	481	2	O6RUW2	O6RUW2 xenopus lae
20	198	40.9	2211	1	FA5 BOVIN	Q28107 bos taurus
21	197	40.7	384	2	O8C8K0	O8C8K0 mus musculu
22	197	40.7	470	2	O8C4U8	O8C4U8 mus musculu
23	197	40.7	480	1	ED13 MOUSE	O35474 mus musculu
24	197	40.7	480	2	O8CBF7	O8CBF7 mus musculu
25	196	40.5	480	1	ED13 HUMAN	O43854 homo sapien
26	196	40.5	480	2	O8N6I0	O8N6I0 homo sapien
27	196	40.5	1460	2	O75ZN0	O75ZN0 pseudonaja
28	194	40.1	407	2	O8AYE0	O8AYE0 brachydanio
29	194	40.1	2119	2	O90X47	O90X47 brachydanio
30	191.5	39.6	858	2	O76470	O76470 lytechinus
31	190	39.3	475	2	O68EW5	O68EW5 xenopus lae

[15]
RP VARIANT HEMA GLN-2228.
RX MEDLINE=89194889; PubMed=2833855;
RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
RA Kazanian H.H.;
RT "Nonsense and missense mutations in hemophilia A: estimate of the
RT relative mutation rate at CG dinucleotides";
RL Am. J. Hum. Genet. 42:718-725(1988).
RN [16]
RP VARIANT HEMA GLY-291.
RX MEDLINE=88220354; PubMed=2835904;
RA Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazanian H.H. Jr.,
RA Antonarakis S.E.;
RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
RT in exon 7 of the factor VIII gene";
RL Am. J. Hum. Genet. 42:867-871(1988).
RN [17]
RP VARIANT HEMA CYS-1708.
RX MEDLINE=89274393; PubMed=2499363;
RA O'Brien D.P., Tuddenham E.G.;
RT "Purification and characterization of factor VIII 1,689-Cys: a
RT nonfunctional cofactor occurring in a patient with severe hemophilia
RT A";
RL Blood 73:2117-2122(1989).
RN [18]
RP VARIANT HEMA CYS-391.
RX MEDLINE=90001543; PubMed=2506948;
RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
RT "An arginine to cysteine amino acid substitution at a critical
RT thrombin cleavage site in a dysfunctional factor VIII molecule";
RL Blood 74:1612-1617(1989).
RN [19]
RP VARIANT HEMA LEU-189.
RX MEDLINE=90057680; PubMed=25110835;
RA Chan V., Chan T.K., Tong T.M., Todd D.;
RT "A novel missense mutation in exon 4 of the factor VIII:C gene
RT resulting in moderately severe hemophilia A";
RL Blood 74:2688-2691(1989).
RN [20]
RP VARIANT HEMA LEU-2326.
RX MEDLINE=89197216; PubMed=2495245;
RA Inaba H., Fujimaki M., Kazanian H.H. Jr., Antonarakis S.E.;
RT "Wild hemophilia A resulting from Arg-to-Leu substitution in exon 26
RT of the factor VIII gene";
RL Hum. Genet. 81:335-338(1989).
RN [21]
RP VARIANT HEMA HIS-391.
RX MEDLINE=89264602; PubMed=2498882;
RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazanian H.H. Jr.,
RA Fujimaki M., Hoyer L.W.;
RT "Direct characterization of factor VIII in plasma: detection of a
RT mutation altering a thrombin cleavage site (arginine-
RT 372-->histidine)";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
RN [22]
RP VARIANT HEMA CYS-1708.
RX MEDLINE=90105723; PubMed=2104766;
RA Arai M., Higuchi M., Antonarakis S.E., Kazanian H.H. Jr.,
RA Phillips J.A. III, Jancso R.L., Hoyer L.W.;
RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
RT Cys) in the factor VIII gene of two unrelated patients with cross-
RT reacting material-positive hemophilia A";
RL Blood 75:384-389(1990).
RN [23]
RP VARIANTS HEMA GLN-2228 AND LEU-2326.
RX MEDLINE=90123183; PubMed=2105106;
RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
RA Mancuso G., Morfini M., de Biasi R., Bardo F., Carbonara A.;
RT "Recurrent mutations and three novel rearrangements in the factor VIII
RT gene of hemophilia A patients of Italian descent";
RL Blood 75:662-670(1990).
RN [24]
RP VARIANT HEMA CYS-391.

RX MEDLINE=90329422; PubMed=1973901;
RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
RT "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
RT internal heavy chain thrombin cleavage site";
RL Br. J. Haematol. 75:73-77(1990).
RN [25]
RP VARIANTS HEMA PHE-1699 AND CYS-1708.
RX MEDLINE=90152691; PubMed=2105906;
RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
RA Kazanian H.H., Antonarakis S.E.;
RT "Characterization of mutations in the factor VIII gene by direct
RT sequencing of amplified genomic DNA";
Query Match 100.0%; Score 484; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 7.3e-44;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPKWLQVDFQKTMKVTGTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
Db 2242 VNNPKWLQVDFQKTMKVTGTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 2301
Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRHPQSW 91
Db 2302 FQGNQDSFTPVVNSLDPPLLTRYLRHPQSW 2332
RESULT 3
ID O18806 PRELIMINARY; PRT: 2343 AA.
AC O18806;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteraia; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Cameron C., Nottley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA Giles A., Lillicrap D.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; AF016234; AAB87412.1; -.
DR HSSP; P00451; 1D7P.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR Pfam; PF00394; Cu-oxidase_1.
DR Pfam; PF00754; F5_P8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00222; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;
Query Match 86.2%; Score 417; DB 2; Length 2343;
Best Local Similarity 84.4%; Pred. No. 1.7e-36;
Matches 76; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
Qy 2 NNPKEWLQVDFQKTMKVTGTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 61
Db 2235 NNPKEWLQVDFQKTMKVTGTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 2294
Qy 62 QGNQDSFTPVVNSLDPPLLTRYLRHPQSW 91
Db 2295 QGNQDSFTPVVNSLDPPLLTRYLRHPQSW 2324

```

RESULT 4
ID 062730 PRELIMINARY; PRT; 2343 AA.
AC 062730; 1998 (TREMELrel. 07, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC EMBL; AF049489; AAC05384.1; -.
CC HSP; P00451; I10D.
CC GO; GO:0005507; F: copper ion binding; IEA.
CC GO; GO:0007155; F: cell adhesion; IEA.
CC InterPro; IPR001117; Cu-oxidase; IEA.
CC InterPro; IPR008972; Cupredoxin.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC SMART; SM00231; FA58C_2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00022; FA58C_3; 2.
CC PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW Signal; Sulfation.
FT CHAIN 1 19
FT DOMAIN 20 2319
FT DOMAIN 20 349
FT DOMAIN 20 199
FT DOMAIN 207 349
FT DOMAIN 399 730
FT DOMAIN 399 573
FT DOMAIN 583 730
FT DOMAIN 760 1640
FT DOMAIN 1683 2008
FT DOMAIN 1683 1845
FT DOMAIN 1855 2008
FT DOMAIN 2008 2156
FT DOMAIN 2161 2313
FT SITE 391 392
FT SITE 1324 1325
FT SITE 1640 1641
FT SITE 1678 1679
FT MOD_RES 367 367
FT MOD_RES 737 737
FT MOD_RES 738 738
FT MOD_RES 742 742
FT MOD_RES 1669 1669
FT MOD_RES 1687 1687
FT DISULFID 173 199
FT DISULFID 547 573
FT DISULFID 1819 1845
FT DISULFID 2008 2156
FT DISULFID 2161 2313
FT CARBOHYD 61 61
FT CARBOHYD 233 233
FT CARBOHYD 259 259
FT CARBOHYD 423 423
FT CARBOHYD 601 601
FT CARBOHYD 880 880
FT CARBOHYD 958 958
FT CARBOHYD 1015 1015
FT CARBOHYD 1022 1022
FT CARBOHYD 1026 1026
FT CARBOHYD 1044 1044
FT CARBOHYD 1076 1076
FT CARBOHYD 1087 1087

Query Match 85.1%; Score 412; DB 2; Length 2343;
Best Local Similarity 83.3%; Pred.No. 6.2e-36;
Matches 75; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 2 NNPKQLVDFDKTKWVGVTTQGVKSLTSMVYKFLISSQDGHQWTLFFQNGKVKVF 61
Db 2235 NNPKQLVDFDKTKWVGVTTQGVKSLTSMVYKFLISSQDGHQWTLFFQNDKVKVF 2294
Qy 62 QGNQDSFTVNSLDPILLTYLRIHQSW 91
Db 2295 QGNRDSSTPVRNLEPPLVARYVRLHFQSW 2324

RESULT 5
FA8 MOUSE STANDARD; PRT; 2319 AA.
AC Q06194;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
CN Name=F8; Synonyms=Cf8, F8c;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=93300511; PubMed=8314577;
RA Elder B., Lakich D., Gitechier J.;
RT "Sequence of the murine factor VIII cDNA.";
RL Genomics 16:374-379(1993).
CC -1- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC as a cofactor for factor IXa when it converts factor X to the
CC activated form, factor Xa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Found in most tissues.

```

```

-1- PTM: The binding of Von Willibrand Factor (VWF) and activation
depend on the sulfation of Tyr-1669.
-1- SIMILARITY: Belongs to the multicopper oxidase family.
-1- SIMILARITY: Contains 3 F5/8 type A domains.
-1- SIMILARITY: Contains 2 F5/8 type C domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; L05573; AAA37395.1; -.
PIR; A47004; A47004.
HSP; P00451; I1D7F.
MGD; MGI:88383; F8.
InterPro; IPR001117; Cu-oxidase.
InterPro; IPR008972; Cupredoxin.
InterPro; IPR000421; FA58_C.
Pfam; PF00394; Cu-oxidase; 1.
Pfam; PF00754; F5_F8_type_C; 2.
SMART; SM00231; FA58C_2.
PROSITE; PS01285; FA58C_1; 2.
PROSITE; PS01286; FA58C_2; 2.
PROSITE; PS00022; FA58C_3; 2.
PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW Signal; Sulfation.
FT CHAIN 1 19
FT DOMAIN 20 2319
FT DOMAIN 20 349
FT DOMAIN 20 199
FT DOMAIN 207 349
FT DOMAIN 399 730
FT DOMAIN 399 573
FT DOMAIN 583 730
FT DOMAIN 760 1640
FT DOMAIN 1683 2008
FT DOMAIN 1683 1845
FT DOMAIN 1855 2008
FT DOMAIN 2008 2156
FT DOMAIN 2161 2313
FT SITE 391 392
FT SITE 1324 1325
FT SITE 1640 1641
FT SITE 1678 1679
FT MOD_RES 367 367
FT MOD_RES 737 737
FT MOD_RES 738 738
FT MOD_RES 742 742
FT MOD_RES 1669 1669
FT MOD_RES 1687 1687
FT DISULFID 173 199
FT DISULFID 547 573
FT DISULFID 1819 1845
FT DISULFID 2008 2156
FT DISULFID 2161 2313
FT CARBOHYD 61 61
FT CARBOHYD 233 233
FT CARBOHYD 259 259
FT CARBOHYD 423 423
FT CARBOHYD 601 601
FT CARBOHYD 880 880
FT CARBOHYD 958 958
FT CARBOHYD 1015 1015
FT CARBOHYD 1022 1022
FT CARBOHYD 1026 1026
FT CARBOHYD 1044 1044
FT CARBOHYD 1076 1076
FT CARBOHYD 1087 1087

```

FT CARBOHYD 1136 1136 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1161 1161 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1192 1192 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1255 1255 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1268 1268 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1273 1273 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1274 1274 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1302 1302 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1316 1316 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1340 1340 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1378 1378 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1797 1797 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 2105 2105 N-linked (GlcNAc. . .) (Potential).
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;
 Query Match 83.9%; Score 406; DB 1; Length 2319;
 Best Local Similarity 82.4%; Pred. No. 2.8e-35;
 Matches 75; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 Qy 1 VNNPKWLVDFQTKMTKVTGTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
 Db 2210 VNDPKQVLQDLQTKMTKVTGTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 2269
 Qy 61 FQGNQDSFTPVNSLDPLLRILRIHPQSW 91
 Db 2270 FQGNQDSFTPVNSLDPLLRILRIHPQSW 2300
 RESULT 6
 ID_FAB_PIG STANDARD; PRT; 2133 AA.
 AC P12263; Q95243;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Coagulation factor VIII precursor (Procoagulant component).
 GN Name=F8; Synonyms=CF8;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Healey J.P., Lubin I.M., Lollar P.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE OF 705-1573 FROM N.A.
 RP MEDLINE=86287369; PubMed=3016730;
 RX Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
 RA Kaufman R.J.;
 RL "A large region (approximately equal to 95 kDa) of human factor VIII is dispensable for in vitro procoagulant activity";
 Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
 [3]
 RN SEQUENCE OF 392-759 FROM N.A.
 RP MEDLINE=94179260; PubMed=7510693;
 RA Lubin I.M., Healey J.P., Scandella D., Runge M.S., Lollar P.;
 RL "Elimination of a major inhibitor epitope in factor VIII";
 J. Biol. Chem. 269:8639-8641(1994).
 CC -1- FUNCTION: Factor VIII, along with calcium and phospholipid, acts as a cofactor for factor IXa when it converts factor X to the activated form, factor Xa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: Belongs to the multicopper oxidase family.
 CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch).
 EMBL; U49517; AAB06705.1; -;
 DR PIR; A25945; A25945.
 DR PIR; T42763; T42763.
 DR HSSP; P00451; 1D7P.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR000421; FA58_C.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF00394; Cu-oxidase; 1.
 DR Pfam; PF00754; F5_F8_type_C; 2.
 DR SMART; SM00231; FA58C_2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR PROSITE; PS00022; FA58C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
 KW Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
 Signal; Sulfation.
 FT SIGNAL 1 19 Potential.
 FT CHAIN 20 2133 Coagulation factor VIII.
 FT DOMAIN 20 357 F5/8 type A 1.
 FT DOMAIN 20 199 Plastocyanin-like 1.
 FT DOMAIN 207 357 Plastocyanin-like 2.
 FT DOMAIN 399 730 F5/8 type A 2.
 FT DOMAIN 399 573 Plastocyanin-like 3.
 FT DOMAIN 583 730 Plastocyanin-like 4.
 FT DOMAIN 760 1599 B.
 FT DOMAIN 1495 1822 F5/8 type A 3.
 FT DOMAIN 1495 1659 Plastocyanin-like 5.
 FT DOMAIN 1669 1822 Plastocyanin-like 6.
 FT DOMAIN 1822 1970 F5/8 type C 1.
 FT DOMAIN 1975 2127 F5/8 type C 2.
 FT SITE 391 392 Cleavage (by thrombin) (By similarity).
 FT SITE 759 760 Cleavage (by thrombin) (By similarity).
 FT SITE 1449 1450 Cleavage (activation) (By similarity).
 FT SITE 1490 1491 Cleavage (by thrombin) (By similarity).
 FT MOD_RES 737 737 Sulfotyrosine (By similarity).
 FT MOD_RES 738 738 Sulfotyrosine (By similarity).
 FT MOD_RES 742 742 Sulfotyrosine (By similarity).
 FT DISULFID 173 199 Probable.
 FT DISULFID 547 573 Probable.
 FT DISULFID 1633 1659 Probable.
 FT DISULFID 1822 1970 By similarity.
 FT DISULFID 1975 2127 By similarity.
 FT CARBOHYD 233 233 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 259 259 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 601 601 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 929 929 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 985 985 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1025 1025 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1111 1111 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1181 1181 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1208 1208 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1245 1245 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1265 1265 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1335 1335 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1408 1408 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1611 1611 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 1919 1919 N-linked (GlcNAc. . .) (Potential).
 FT CONFLICT 713 713 N -> M (in Ref. 2).
 FT CONFLICT 734 734 I -> T (in Ref. 2).
 FT CONFLICT 792 792 G -> Q (in Ref. 2).
 FT CONFLICT 1133 1133 E -> F (in Ref. 2).
 FT CONFLICT 1191 1191 I -> L (in Ref. 2).
 FT CONFLICT 1209 1209 R -> F (in Ref. 2).
 FT CONFLICT 1437 1437 C -> G (in Ref. 2).
 FT CONFLICT 1456 1456 F -> R (in Ref. 2).
 FT CONFLICT 1539 1539 F -> R (in Ref. 2).
 FT CONFLICT 1546 1546 Q -> N (in Ref. 2).
 SQ SEQUENCE 2133 AA; 239304 MW; 152BBAB997F570DA CRC64;
 Query Match 81.0%; Score 392; DB 1; Length 2133;

Best Local Similarity 79.1%; Pred. No. 8.8e-34;
Matches 72; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGVTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
DB 2024 VSSAEWLVQDLQKTVKVTGVTGGVKSLLTSMYKFLISSQDGRWTLFLQDGHQTKV 2083

QY 61 FQGNQDSFTPVVNSLDPPLRLYRIHPQSW 91
DB 2084 FQGNQDSFTPVVNSLDPPLRLYRIHPQSW 2114

RESULT 7
Q7TN96 PRELIMINARY; PRT; 2258 AA.
AC Q7TN96;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Factor VIII.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; Tissue=Liver;
RA Watzka M., Geisen C., Seifried E., Oldenburg J.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
DB EMBL; AY362193; AAQ21580.1; -.
DR HSP; P00451; ICFG.
DR GO; GO:0005507; F: copper ion binding; IEA.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C_2; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 1.
DR PROSITE; PS50022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
FT NON TER 1
SQ SEQUENCE 1377 AA; 152728 MW; 15ACFD7F765665CD CRC64;

Query Match 60.1%; Score 291; DB 2; Length 1377;
Best Local Similarity 55.6%; Pred. No. 7.1e-23;
Matches 50; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 2 NNPKEWLQVDFQKTMKVTGVTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 61
DB 1369 NSPSEWLQVDFQKTMKVTGVTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 1328

QY 62 QGNQDSFTPVVNSLDPPLRLYRIHPQSW 91
DB 1329 RANRDTSTVINSLEPPLFARYVRIHPRH 1358

RESULT 9
Q804W6 PRELIMINARY; PRT; 1639 AA.
AC Q804W6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coagulation factor VIII.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]

RP SEQUENCE FROM N.A.
RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA Tuddenham E.G.D., McVey J.H.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
DB EMBL; AF465279; AAQ33374.1; -.
DR HSP; P00451; ICFG.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C_2; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS50022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN 2.
SQ SEQUENCE 1639 AA; 185203 MW; 881778AAFF809006 CRC64;

Query Match 49.6%; Score 240; DB 2; Length 1639;
Best Local Similarity 48.9%; Pred. No. 3.5e-17;
Matches 45; Conservative 12; Mismatches 33; Indels 2; Gaps 1;

Best Local Similarity 72.5%; Score 351; DB 2; Length 2258;
Matches 63; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGVTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
DB 2149 VNDPTQWLQDLQRTVKTGVTGGVKSLLTSMYKFLISSQDGRWTLFLQDGHQTKV 2208

QY 61 FQGNQDSFTPVVNSLDPPLRLYRIHPQSW 91
DB 2209 FQGNQDSFTPVVNSLDPPLRLYRIHPQSW 2239

RESULT 8
Q804X3 PRELIMINARY; PRT; 1377 AA.
AC Q804X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coagulation factor VIII (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

141	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6J; TISSUE=Liver;
RC	MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX	
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA	Kanno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to
RT	

RT "Molecular cloning and characterization of P47, a novel boar sperm-associated zona pellucida-binding protein homologous to a family of mammalian secretory proteins.";
 RT Biol. Reprod. 58:1057-1064 (1998).
 RN (4)
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), TISSUE SPECIFICITY,
 RP DEVELOPMENTAL STAGE, INDUCTION, AND GLYCOSYLATION.
 RC STRAIN=BAUB/C; TISSUE=Mammary gland;
 RX MEDLINE=99120894; PubMed=9920772; DOI=10.1006/bbrc.1998.0107;
 RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
 RT "Lactation-dependent expression of an mRNA splice variant with an exon
 for a multiply O-glycosylated domain of mouse milk fat globule
 glycoprotein MFG-E8.";
 RL Biochem. Biophys. Res. Commun. 254:522-528 (1999).
 RN (5)
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP STRAIN=NOD; TISSUE=Dendritic cell;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nkaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gutcinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yangdisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai K., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yaeunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN (6)
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Heisler F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshitoki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: Specific ligand for the alpha-v/beta-3 and alpha-v/beta-
 5 receptors. Also binds to phosphatidylserine-enriched cell

CC surfaces in a receptor-independent manner. Zona pellucida-binding
 CC protein which may play a role in gamete interaction (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Long;
 CC IsoId=P21956-1; Sequence=Displayed;
 CC Name=2; Synonyms=Short;
 CC IsoId=P21956-2; Sequence=VSP 009880;
 CC -1- TISSUE SPECIFICITY: Mammary epithelial cell surfaces and
 CC spermatozoan. Isoform 2 is present in brain, heart, kidney and
 CC spleen and at low levels in lung, liver, small intestine and
 CC testis.
 CC -1- DEVELOPMENTAL STAGE: Isoform 1 and isoform 2 are detectable in
 CC mammary tissue from non-pregnant animals, with isoform 2 being
 CC predominant. Levels of isoform 1 increase during gestation and
 CC lactation while levels of isoform 2 decrease.
 CC -1- INDUCTION: Isoform 1 is induced by insulin, prolactin and
 CC hydrocortisone in mammary epithelial cells. Expression of isoform
 CC 2 is repressed by the same treatment.
 CC -1- PTM: N-glycosylated. Isoform 1 also exists in both an O-
 CC glycosylated and a non-O-glycosylated form.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M38337; AAA39534.1; -;
 CC EMBL; Y11684; CAA72380.2; -;
 CC EMBL; AB021130; BAA35180.1; -;
 CC EMBL; AB025280; BAA76386.1; -;
 CC EMBL; AK089211; BAC40794.1; -;
 CC EMBL; BC003892; AAH03892.1; -;
 CC EMBL; BC003904; AAH03904.1; -;
 CC PIR; A36479; A36479.
 CC HSP; P00740; IEDM.
 CC MGD; MGI:102768; Mige8.
 CC GO; GO:0009897; C:external side of plasma membrane; IDA.
 CC GO; GO:0005615; C:extracellular space; IDA.
 CC GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
 CC GO; GO:0005178; F:integrin binding; IDA.
 CC GO; GO:0008429; F:phosphatidylethanolamine binding; IDA.
 CC GO; GO:0001786; F:phosphatidylserine binding; IDA.
 CC GO; GO:0006911; P:phagocytosis, binding; IDA.
 CC GO; GO:0005766; P:phagocytosis, engulfment; IDA.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000421; FA58_C.
 CC InterPro; IPR008979; GalBind_like.
 CC InterPro; IPR006210; IEGF.
 CC Pfam; PF00008; EGF; 2.
 CC Pfam; PF00754; F5_F8_type_C; 2.
 CC PRINTS; PR00010; EGFBL00D.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 2.
 CC PROSITE; PS00026; EGF_3; 2.
 CC PROSITE; PS01285; FA58C_1; 2.
 CC PROSITE; PS01286; FA58C_2; 2.
 CC PROSITE; PS00022; FA58C_3; 2.
 CC PROSITE; PS01286; FA58C_2; 2.
 CC KW Alternative splicing; Cell adhesion; Direct protein sequencing;
 KW EGF-like domain; Fertilization; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 463 Lactadherin.
 FT DOMAIN 24 61 EGF-like 1.

```
FT DOMAIN 64 108 EGF-like 2.
FT DOMAIN 148 303 F5/8 type C 1.
FT DOMAIN 308 463 F5/8 type C 2.
FT SITE 87 89 Cell attachment site (Potential).
FT DISULFID 28 39 By similarity.
FT DISULFID 33 49 By similarity.
FT DISULFID 51 60 By similarity.
FT DISULFID 58 79 By similarity.
FT DISULFID 73 96 By similarity.
FT DISULFID 98 107 By similarity.
FT DISULFID 148 303 By similarity.
FT DISULFID 290 294 By similarity.
FT DISULFID 308 463 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 266 266 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 316 316 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 426 426 N-linked (GlcNAc...) (Potential).
FT VARSPIC 110 147 ETNYNLDGEYMTTAVPTAVPTAPTPDLSNNLASR ->
FT G (in isoform 2).
FT /FTid=VSP 009880.
FT N -> D (in Ref. 1; AA sequence).
FT CONFLICT 35 35
Query Match 42.4%; Score 205; DB 1; Length 463;
Best Local Similarity 42.2%; Pred. No. 6e-14;
Matches 38; Conservative 16; Mismatches 36; Indels 0; Gaps 0;
QY 2 NNPKEWLQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV 61
DB 361 NSAKWLQVGLGTQRTQVTGTLTQARDFGHQIVASIKVAHSDGVDGQWTVTEGGSSKVF 420
QY 62 QNQNDSFTPVNSLDPPLLYRIHPQSW 91
DB 421 QGNLDNNSHKKNIFEKPFMARYRVLFVSW 450
RESULT 14
Q7TPK2 ID Q7TPK2 PRELIMINARY; PRT; 2102 AA.
AC Q7TPK2;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Ac2-120.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1];
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Yang K.J., Yan H.M., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; AY321333; AAP86265.1; -.
DR HSP; P12259; ICZT.
DR GO; GO:007596; P: blood coagulation; IEA.
DR GO; GO:007155; P: cell adhesion; IEA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR009271; LSPD-.
DR Pfam; PF00754; F5 F8 type_C; 2.
DR Pfam; PF06049; LSPR; 20.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 1.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
SQ SEQUENCE 2102 AA; 237257 MW; 001BCB7DE31AC560 CRC64;
Query Match 42.1%; Score 204; DB 2; Length 2102;
Best Local Similarity 41.3%; Pred. No. 4.3e-13;
Matches 38; Conservative 20; Mismatches 32; Indels 2; Gaps 1;
```

```
QY 2 NNPKEWLQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV--K 59
DB 1995 NNNKQWLQIDLLKIKKVTATVTCCKSLSEMVKSYSLYSDQGVSWKPYRKSSMVDK 2054
QY 60 VFQGNDSFTPVNSLDPPLLYRIHPQSW 91
DB 2055 IFEGNSNTKGHKNFNPPPIISRFIRIPKTV 2086
RESULT 15
Q804X4 ID Q804X4 PRELIMINARY; PRT; 745 AA.
AC Q804X4;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Coagulation factor V (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1];
RP SEQUENCE FROM N.A.
RA Davidson C.J., Hirt P., Lal K., Snell P., Elgar G.,
RA Tuddenham E.G.D., McVey J.H.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; AF465271; AAC33366.1; -.
DR HSP; P12259; ICZT.
DR GO; GO:007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR008979; Gal Bind like.
DR Pfam; PF00754; F5 F8 type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
DR NON TER 1
SQ SEQUENCE 745 AA; 85626 MW; A264587A348C29EE CRC64;
Query Match 41.9%; Score 203; DB 2; Length 745;
Best Local Similarity 40.2%; Pred. No. 1.7e-13;
Matches 37; Conservative 22; Mismatches 31; Indels 2; Gaps 1;
QY 2 NNPKEWLQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV--VK 59
DB 636 NNNKQWLQIDLLKIKKVTATVTCCKSLSEMVKSYSLYSDQGVSWKPYRKSSMVDK 695
QY 60 VFQGNDSFTPVNSLDPPLLYRIHPQSW 91
DB 696 VFSGNDSRGHVKHFFNPPPIISRFIRIPKTV 727
Search completed: April 20, 2005, 17:21:04
Job time : 23.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:08:58 ; Search time 97.6188 Seconds
(without alignments)
1240.089 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332

Perfect score: 1666

Sequence: 1 LIGHLHAGMSTFLVYSNK.....VNSLDPLLRLLYLRHPSW 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	684	2	AAR73022 Human Fac
2	1666	100.0	684	2	AAR74091 Factor-VI
3	1666	100.0	770	3	ADP20767 Factor VI
4	1666	100.0	790	5	ADP64594 Recombina
5	1666	100.0	1014	1	AAP71139 Factor VII
6	1666	100.0	1383	2	AAW33227 Procoagul
7	1666	100.0	1383	2	AAW33228 Procoagul
8	1666	100.0	1383	2	AAW33229 Procoagul
9	1666	100.0	1421	8	ADQ37613 Human Fac
10	1666	100.0	1421	8	ADQ37598 Human Fac
11	1666	100.0	1422	8	ADQ37591 Human Fac
12	1666	100.0	1422	8	ADQ37606 Human Fac
13	1666	100.0	1424	1	AAP80268 Modified
14	1666	100.0	1424	1	AAP91169 Sequence
15	1666	100.0	1424	4	AAB48842 Mutant ma
16	1666	100.0	1424	5	AAO18622 Human mat
17	1666	100.0	1425	1	AAP80267 Modified
18	1666	100.0	1425	8	ADQ37605 Human Fac
19	1666	100.0	1425	8	ADQ37590 Human Fac
20	1666	100.0	1428	8	ADQ37599 Human Fac
21	1666	100.0	1428	8	ADQ37604 Human Fac
22	1666	100.0	1428	8	ADQ37589 Human Fac
23	1666	100.0	1428	8	ADQ37614 Human Fac
24	1666	100.0	1434	8	ADQ37594 Human Fac
25	1666	100.0	1434	8	ADQ37609 Human Fac

26	1666	100.0	1435	8	ADQ37600 Human Fac
27	1666	100.0	1435	8	ADQ37615 Human Fac
28	1666	100.0	1437	8	ADQ37593 Human Fac
29	1666	100.0	1437	8	ADQ37608 Human Fac
30	1666	100.0	1438	3	AAO1262 B-domain
31	1666	100.0	1438	7	ADF31316 Variant o
32	1666	100.0	1438	7	ADM75988 Modified
33	1666	100.0	1438	7	ADM75986 Modified
34	1666	100.0	1438	7	ADM75989 Modified
35	1666	100.0	1438	7	ADM75983 Wild-type
36	1666	100.0	1440	2	AAR12971 Factor VI
37	1666	100.0	1440	8	ADQ37607 Human Fac
38	1666	100.0	1440	8	ADQ37592 Human Fac
39	1666	100.0	1441	8	ADQ37612 Human Fac
40	1666	100.0	1441	8	ADQ37597 Human Fac
41	1666	100.0	1442	8	ADQ37616 Human Fac
42	1666	100.0	1442	8	ADQ37601 Human Fac
43	1666	100.0	1444	8	ADQ37611 Human Fac
44	1666	100.0	1444	8	ADQ37596 Human Fac
45	1666	100.0	1447	5	ABG92541 5Arg B-do

ALIGNMENTS

RESULT 1

AAR73022
ID AAR73022 standard; peptide; 684 AA.

AC AAR73022;

XX 25-MAR-2003 (revised)

DT 21-NOV-1995 (first entry)

XX Human Factor-VIII C-terminal fragment.

XX Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.

OS Homo sapiens.

PN WO9513301-A1.

XX 18-MAY-1995.

PF 10-NOV-1994; 94WO-DK000424.

XX 12-NOV-1993; 93DK-00001281.

PA (NOVO) NOVO-NORDISK AS.

PI Persson E;

DR WPI; 1995-194038/25.

XX Crosslinked Factor VIII polypeptide which is stable - is prep'd. using bis(sulphosuccinimidyl) substrate or disuccinimidyl substrate in the presence of polysorbate 80 to produce a coagulant with long lasting activity.

PS Disclosure; Page 24; 36pp; English.

XX This is the C-terminal fragment (corresponding to AAs 1649-2332) of human Factor-VIII which may be crosslinked resulting in increased stability and retention of high activity over extended periods of time after activation by thrombin. The polypeptide is used to prevent or treat diseases caused by the absence or deficiency of Factor-VIII in a subject such as haemophilia. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 684 AA;

Query Match 100.0%; Score 1666; DB 2; Length 684;

Best Local Similarity 100.0%; Pred. No. 2.3e-168;

Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYSNKQCTPLGWSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 60
 DB 353 LIGELHAGMSTFLVYSNKQCTPLGWSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 412
 QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGAROKFESSLYISQFIIMYSLDGKKWQTYRGNST 120
 DB 413 AMSTKEPFSWIKVDLLAPMIHGIKTQGAROKFESSLYISQFIIMYSLDGKKWQTYRGNST 472
 QY 121 GTLMVFFGNVDSGGIKHNIENPPIIARYIRLHPHYSIRSTLRMELMCGDLNCSMPGLM 180
 DB 473 GTLMVFFGNVDSGGIKHNIENPPIIARYIRLHPHYSIRSTLRMELMCGDLNCSMPGLM 532
 QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKFWLQVDFQKTMKV 240
 DB 533 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKFWLQVDFQKTMKV 592
 QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
 DB 593 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 652
 QY 301 LLTRYLRHQPQSW 313
 DB 653 LLTRYLRHQPQSW 665

RESULT 2
 AAR74091
 ID AAR74091 standard; protein; 684 AA.
 AC AAR74091;
 XX 25-MAR-2003 (revised)
 DT 04-NOV-1995 (first entry)
 XX Factor-VIII light chain C-terminal fragment.
 XX human; Factor VIII; light chain; C-terminal fragment; thrombin cleavage;
 KW blood-clotting.
 XX Homo sapiens.
 XX WO9513300-A1.
 XX 18-MAY-1995.
 XX 10-NOV-1994; 94WO-DK000423.
 XX 12-NOV-1993; 93DK-00001280.
 XX (NOVO) NOVO-NORDISK AS.
 PA Kjalke M, Ezban Rasmussen M;
 PI WPI; 1995-194037/25.
 XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
 PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
 PT activity, may be used to treat patients who have developed antibodies to
 PT C-terminal epitope(s) of Factor VIII.

PS Claim 4; Page 33-35; 51pp; English.
 XX The sequence represents C-terminal residues 1649-2332 of a human Factor-
 CC VIII light chain. The sequence is produced by treating a full-length
 CC Factor-VIII polypeptide with a protease, e.g. thrombin. The fragment may
 CC be produced recombinantly in conjunction with a C-terminally truncated
 CC heavy chain to reduce production costs and improve safety, and production
 CC levels and stability are higher than for the full-length form. When the
 CC fragment is used along with a C-terminally truncated heavy chain, it may
 CC be used to treat patients who have developed antibodies against epitopes
 CC in the C-terminal part of the heavy chain. (Updated on 25-MAR-2003 to
 CC correct PN field.)

XX SQ Sequence 684 AA;
 Query Match 100.0%; Score 1666; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 2.3e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LIGELHAGMSTFLVYSNKQCTPLGWSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 60
 DB 353 LIGELHAGMSTFLVYSNKQCTPLGWSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 412
 QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGAROKFESSLYISQFIIMYSLDGKKWQTYRGNST 120
 DB 413 AMSTKEPFSWIKVDLLAPMIHGIKTQGAROKFESSLYISQFIIMYSLDGKKWQTYRGNST 472
 QY 121 GTLMVFFGNVDSGGIKHNIENPPIIARYIRLHPHYSIRSTLRMELMCGDLNCSMPGLM 180
 DB 473 GTLMVFFGNVDSGGIKHNIENPPIIARYIRLHPHYSIRSTLRMELMCGDLNCSMPGLM 532
 QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKFWLQVDFQKTMKV 240
 DB 533 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKFWLQVDFQKTMKV 592
 QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
 DB 593 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 652
 QY 301 LLTRYLRHQPQSW 313
 DB 653 LLTRYLRHQPQSW 665

RESULT 3
 ADP20767
 ID ADP20767 standard; protein; 770 AA.
 AC ADP20767;
 XX 12-AUG-2004 (first entry)
 DT Factor VIII protein.
 DE A-domain; C-domain; blood coagulation; factor VIII;
 KW von Willebrand factor.
 XX Unidentified.
 OS KR99066381-A.
 PN 16-AUG-1999.
 PD 24-JAN-1998; 98KR-00002264.
 PF 24-JAN-1998; 98KR-00002264.
 PR (GREG) KOREA GREEN CROSS CORP.
 PA Kim HC, Byun TH, Kim SM, Kim JM, Lee SG;
 PI WPI; 2000-547435/50.
 DR N-PSDB; ADP20766.
 XX Cell line expressing A-domain, C-domain of blood coagulation factor VIII
 PT and von Willebrand factor simultaneously, method - for manufacturing
 PT recombinant blood coagulation factor VIII using the same.
 XX Disclosure; Fig 2; 21pp; Korean.
 CC The invention relates to a novel cell line expressing A-domain, C-domain
 CC of blood coagulation factor VIII and von Willebrand factor
 CC simultaneously, and a method for manufacturing recombinant blood
 CC coagulation factor VIII using the same. The present sequence represents a
 CC factor VIII protein.

XX SQ Sequence 770 AA;
 Query Match 100.0%; Score 1666; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 2.8e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYNSKQCTPLGMSGHIRDFQITASGOYGOWAPKLARLHYSGSIN 60
 DB 439 LIGELHAGMSTFLVYNSKQCTPLGMSGHIRDFQITASGOYGOWAPKLARLHYSGSIN 498
 QY 61 AWSTKEPPSWIKVDLLAPMIHGIKTQAROKFSSLYISQFTIMYSLDGKKWQTYRGNST 120
 DB 499 AWSTKEPPSWIKVDLLAPMIHGIKTQAROKFSSLYISQFTIMYSLDGKKWQTYRGNST 558
 QY 121 GTLMVFFGNVDSGKKNIFNPPPIIARIYRLHPTHYSIRSLRMELMGCDLNSCSMPGLM 180
 DB 559 GTLMVFFGNVDSGKKNIFNPPPIIARIYRLHPTHYSIRSLRMELMGCDLNSCSMPGLM 618
 QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWQVNNPKEWLQVDFOKTMKV 240
 DB 619 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWQVNNPKEWLQVDFOKTMKV 678
 QY 241 TGVTTQGVKSLTSMYKKEFLISSODGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 300
 DB 679 TGVTTQGVKSLTSMYKKEFLISSODGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 738
 QY 301 LLTRYLRIHQSW 313
 DB 739 LLTRYLRIHQSW 751

RESULT 4
 ADE64594
 ID ADE64594 standard; protein; 790 AA.
 AC ADE64594;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Recombinant blood coagulation factor VIII protein, SEQ ID 4.
 XX
 KW blood coagulation factor VIII; type-A haemophilia.
 XX
 OS Unidentified.
 XX
 PN CN1361178-A.
 XX
 PD 31-JUL-2002.
 XX
 PF 29-DEC-2000; 2000CN-00137779.-
 XX
 PR 29-DEC-2000; 2000CN-00137779.
 XX
 PA (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
 XX
 PI Qi Z, Wang Q, Chen C;
 XX
 DR WPI; 2002-741852/81.
 DR N-PSDB; ADE64593.
 XX
 FT New recombinant blood coagulation factor VIII and its production process
 FT and medicinal composition.
 XX
 PS Claim 2; SEQ ID NO 4; 31pp; Chinese.
 XX
 CC The invention relates to a novel recombinant blood coagulation factor
 CC VIII, its production process and its medicinal composite for treating
 CC type-A haemophilia. The invention further comprises a medicinal
 CC composition containing the blood coagulation factor which promotes blood
 CC coagulation to the blood plasma of type-A haemophilia patients. This
 CC sequence represents a recombinant blood coagulation factor VIII protein
 CC of the invention.

XX SQ Sequence 790 AA;
 Query Match 100.0%; Score 1666; DB 5; Length 790;
 Best Local Similarity 100.0%; Pred. No. 2.9e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYNSKQCTPLGMSGHIRDFQITASGOYGOWAPKLARLHYSGSIN 60
 DB 459 LIGELHAGMSTFLVYNSKQCTPLGMSGHIRDFQITASGOYGOWAPKLARLHYSGSIN 518
 QY 61 AWSTKEPPSWIKVDLLAPMIHGIKTQAROKFSSLYISQFTIMYSLDGKKWQTYRGNST 120
 DB 519 AWSTKEPPSWIKVDLLAPMIHGIKTQAROKFSSLYISQFTIMYSLDGKKWQTYRGNST 578
 QY 121 GTLMVFFGNVDSGKKNIFNPPPIIARIYRLHPTHYSIRSLRMELMGCDLNSCSMPGLM 180
 DB 579 GTLMVFFGNVDSGKKNIFNPPPIIARIYRLHPTHYSIRSLRMELMGCDLNSCSMPGLM 638
 QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWQVNNPKEWLQVDFOKTMKV 240
 DB 639 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWQVNNPKEWLQVDFOKTMKV 698
 QY 241 TGVTTQGVKSLTSMYKKEFLISSODGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 300
 DB 699 TGVTTQGVKSLTSMYKKEFLISSODGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 758
 QY 301 LLTRYLRIHQSW 313
 DB 759 LLTRYLRIHQSW 771

RESULT 5
 AAP71139
 ID AAP71139 standard; protein; 1014 AA.
 XX
 AC AAP71139;
 XX
 DT 25-MAR-2003 (revised)
 DT 14-MAY-1991 (first entry)
 XX
 DE Factor VIII:c variant.
 XX
 KW Factor VIII:c; variant; proteolysis; resistance;
 KW pro coagulation activity.
 XX
 OS Homo sapiens
 XX
 PN W08707144-A.
 XX
 PD 03-DEC-1987.
 XX
 PF 29-MAY-1987; 87WO-US001299.
 XX
 PR 29-MAY-1986; 86US-00868410.
 PR 18-NOV-1986; 86US-00932767.
 PR 09-DEC-1986; 86US-00939658.
 XX
 PA (GEM) GENETICS INST INC.
 XX
 PI Kaufman RJ, Pittman D, Toole JJ;
 XX
 DR WPI; 1987-348539/49.
 XX
 FT New deletion and replacement variants of factor 8:c - resistant to
 FT proteolysis but retaining pro-coagulant activity, and new DNA coding
 FT sequences.
 XX
 PS Disclosure; Page 1; 42pp; English.
 XX
 CC The full-length human factor VIII:c cDNA has been set forth in detail in
 CC W08501961. This sequence is an example of the formula A-X-B, wherein
 CC A=Ala1-Arg372, B=Ser1690-Tyr2332 and X=0-1316 amino acids substantially

CC duplicative of sequences of amino acids within the sequence Arg372-Ser1690 of the full-length sequence. Here X=0; producing a Arg371-Ser1690 fusion protein. One or more deletions or replacements at Arg 220, 226, 279, 282, 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce variants which are more resistant to specific proteolytic cleavage compared with natural factor VIII:c. Pro-coagulant activity and thrombin activatability have been retained. See also AAP71726-29. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1014 AA;

Query Match 100.0%; Score 1666; DB 1; Length 1014;
Best Local Similarity 100.0%; Pred. No. 4.2e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYVSNKCOPLGNWASGHIRDFQITASGOYGOWAPKLARLHYSGSIN 60
DB 683 LIGELHAGMSTFLVYVSNKCOPLGNWASGHIRDFQITASGOYGOWAPKLARLHYSGSIN 742
QY 61 AWSTKEPESWIKVDLLAPMIHGIKTQGAOKPSSLYISQFIIMYSLDGKKWQYRGNST 120
DB 743 AWSTKEPESWIKVDLLAPMIHGIKTQGAOKPSSLYISQFIIMYSLDGKKWQYRGNST 802
QY 121 GTLMVFFGNDVSSGIGKNIENPPIIARYIRLHPTHTYIRSLRMLMGCDLNSCSMPLGM 180
DB 803 GTLMVFFGNDVSSGIGKNIENPPIIARYIRLHPTHTYIRSLRMLMGCDLNSCSMPLGM 862
QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKREWLVQDFQKTMKV 240
DB 863 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKREWLVQDFQKTMKV 922
QY 241 TGVTTQGVKSLTSMYKVEFLISSODGHWTLTFQNGKVKVFGQNGDSFTFVNSLDLDP 300
DB 923 TGVTTQGVKSLTSMYKVEFLISSODGHWTLTFQNGKVKVFGQNGDSFTFVNSLDLDP 982
QY 301 LLTRYLRHPQSW 313
DB 983 LLTRYLRHPQSW 995

RESULT 6
AAW33227
ID AAW33227 standard; protein; 1383 AA.
XX AAW33227;
AC AAW33227;
XX AAW33227;
DT 21-OCT-2004 (revised)
DT 30-APR-1998 (first entry)
XX
DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
XX Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
XX Homo sapiens.
OS Synthetic.
XX
FH Key
FH Region Location/Qualifiers
FT 1..346 /note= "factor VIIIA heavy chain"
FT Domain 1..329 /note= "A1 domain"
FT Domain 1..179 /note= "plastocyanin-like domain 1"
FT Disulfide-bond 153..179 /note= "probable"
FT Domain 187..329 /note= "plastocyanin-like domain 2"
FT Cleavage-site 372..373 /note= "by thrombin"

FT Domain 380..711 /note= "A2 domain"
FT Domain 380..554 /note= "plastocyanin-like domain 3"
FT Disulfide-bond 528..554 /note= "probable"
FT Domain 564..711 /note= "plastocyanin-like domain 4"
FT Region 711..746 /note= "a spacer of the sequence SFSQSRHPSRQKFNATIPENDIEKTPWF AHRTPMPKIQNVSSDLLMLL is inserted between domains A2 and A3"
FT Misc-difference 740 /label= R740A
FT Region 741..1383 /note= "wild type Arg replaced with Ala"
FT Domain 746..1073 /note= "factor VIIIA light chain"
FT Domain 1073..1221 /note= "A3 domain"
FT Domain 1226..1378 /note= "C1 domain"
FT Domain 1226..1378 /note= "C2 domain"
XX W09740145-A1.
PN 30-OCT-1997.
XX 24-APR-1997; 97WO-US006563.
XX 24-APR-1996; 96US-0016117P.
XX 15-MAY-1996; 96US-0017785P.
XX (UNMI) UNIV MICHIGAN.
XX Kaufman RJ, Pipe SW, Amano K;
XX WPI; 1997-535830/49.
XX Modified human pro-coagulant active factor VIII - can be administered to haemophiliacs, i.e. factor VIII replacement therapy.
XX Claim 20; Page; 57pp; English.
XX The present sequence represents a novel pro-coagulant active factor VIII (FVIII) mutant protein, comprising a deletion of the B domain and von Willebrand factor binding site, a mutation at Arg740 and an addition of an amino acid sequence spacer between the A2 and A3 domains. Factor VIII, along with calcium and phospholipid, acts as a cofactor for factor IXa, when it converts factor X to the activated form (factor Xa). FVIII is the coagulation factor deficient in the X-chromosome-linked bleeding disorder haemophilia A. Several other mutant FVIII proteins have also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of recombinant secretion at higher levels than typically obtained with wild type FVIII and retains pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein C (APC) cleavage. The present FVIII mutant can form a more stable configuration, and have an approximate 5-fold increase in specific activity compared to purified wild type FVIII, while increasing their binding affinity to von Willebrand factor improves their stability. The FVIII proteins can be administered to haemophiliacs, i.e. FVIII replacement therapy, while the nucleic acid molecule can be used for gene therapy. note: this sequence does not appear in the specification; it was created using sequences from the given references
XX Revised record issued on 21-OCT-2004 : Correction to feature table key
XX Sequence 1383 AA;
XX Query Match 100.0%; Score 1666; DB 2; Length 1383;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-168;
XX Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGELHAGMSTFLVYSNKCQTPLGMSCHIRDFQITASQYQGWAPKRLARLHYSGSIN 60
 Db 1052 LIGELHAGMSTFLVYSNKCQTPLGMSCHIRDFQITASQYQGWAPKRLARLHYSGSIN 1111
 Qy 61 AWSKPEFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
 Db 1112 AWSKPEFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1171
 Qy 121 GTLMVFFGNVDSGIGKHINFPPIIARIYIRLHPHYSIRSTLRMELMGCDLNSCSPMLGM 180
 Db 1172 GTLMVFFGNVDSGIGKHINFPPIIARIYIRLHPHYSIRSTLRMELMGCDLNSCSPMLGM 1231
 Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHOGESNAWRPQVNNPKEWLQVDFQKTKV 240
 Db 1232 ESKAISDAQITASSYFTNMFATWSPSKARLHOGESNAWRPQVNNPKEWLQVDFQKTKV 1291
 Qy 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 300
 Db 1292 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 1351
 Qy 301 LLTRYLRHPOSW 313
 Db 1352 LLTRYLRHPOSW 1364

RESULT 7
 AAW33228
 ID AAW33228 standard; protein; 1383 AA.
 AC AAW33228;
 XX
 XX
 DT 21-OCT-2004 (revised)
 DT 30-APR-1998 (first entry)
 XX
 XX
 DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
 XX
 KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX
 OS Homo sapiens.
 OS Synthetic.

FH Key Location/Qualifiers
 FT Region 1..346
 FT Domain /note= "factor VIIIA heavy chain"
 FT Domain /note= "A1 domain"
 FT Domain /note= "A2 domain"
 FT Disulfide-bond 153..179
 FT Domain /note= "probable"
 FT Domain /note= "probable"
 FT Domain /note= "probable"
 FT Misc-difference 336
 FT /label= R3361
 FT /note= "wild type Arg replaced with Ile"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Domain 380..711
 FT Domain /note= "A2 domain"
 FT Domain 380..554
 FT /note= "plastocyanin-like domain 3"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Misc-difference 562
 FT /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 FT Domain 564..711
 FT /note= "plastocyanin-like domain 4"
 FT Region 711..746

FT /note= "a spacer of the sequence
 FT SFQNSRHPSTRQKFNATTIPENDIEKTDWPF AHRTPMPKIQNVSSDLMML
 FT is inserted between domains A2 and A3"
 FT 740
 FT Misc-difference /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 FT Region 741..1383
 FT Domain /note= "factor VIIIA light chain"
 FT Domain 746..1073
 FT Domain /note= "A3 domain"
 FT Domain 1073..1221
 FT /note= "C1 domain"
 FT Domain 1226..1378
 FT /note= "C2 domain"
 XX
 PN WO9740145-A1.
 XX
 PD 30-OCT-1997.
 XX
 PP 24-APR-1997; 97WO-US0006563.
 PR 24-APR-1996; 96US-0016117P.
 PR 15-MAY-1996; 96US-0017785P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Kaufman RJ, . Pipe SW, Amano K;
 XX WPI, 1997-535830/49.
 DR
 XX Modified human pro-coagulant active factor VIII - can be administered to
 FT haemophiliacs, i.e. factor VIII replacement therapy.
 FT
 PS Claim 18; Page; 57pp; English.
 CC
 CC The present sequence represents a novel pro-coagulant active factor VIII
 CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 CC Willebrand factor binding site, mutations R336I, R562K and R740A and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains.
 CC Factor VIII, along with calcium and phospholipid, acts as a cofactor for
 CC factor IXa, when it converts factor X to the activated form (factor Xa).
 CC FVIII is the coagulation factor deficient in the X-chromosome-linked
 CC bleeding disorder haemophilia A. Several other mutant FVIII proteins have
 CC also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
 CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII mutant
 CC R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein
 CC C (APC) cleavage. The FVIII mutant comprising a deletion of the B domain
 CC and von Willebrand factor binding site, a mutation at Arg740 and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains
 CC can form a more stable configuration, and have an approximate 5-fold
 CC increase in specific activity compared to purified wild type FVIII, while
 CC increasing their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy. note: this sequence does not appear in the
 CC specification; it was created using sequences from the given references
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to feature table key
 XX
 SQ Sequence 1383 AA;
 Query Match 100.0%; Score 1666; DB 2; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 6.7e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LIGELHAGMSTFLVYSNKCQTPLGMSCHIRDFQITASQYQGWAPKRLARLHYSGSIN 60
 Db 1052 LIGELHAGMSTFLVYSNKCQTPLGMSCHIRDFQITASQYQGWAPKRLARLHYSGSIN 1111
 Qy 61 AWSKPEFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
 Db 1112 AWSKPEFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1171

QY 121 GTLMVFFGNVDSSGIKHNIENPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 180
 DB 1172 GTLMVFFGNVDSSGIKHNIENPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 1231
 QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPOVNNPKWLOVDFOKTMKV 240
 DB 1232 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPOVNNPKWLOVDFOKTMKV 1291
 QY 241 TGVTTQGVKSLLTSMYKKEFLISSQDGHQWTLFPQNGKVKVFOGQNDSTFPVNSLDPP 300
 DB 1292 TGVTTQGVKSLLTSMYKKEFLISSQDGHQWTLFPQNGKVKVFOGQNDSTFPVNSLDPP 1351
 QY 301 LLTRYLRHPQSW 313
 DB 1352 LLTRYLRHPQSW 1364
 RESULT 8
 AAW33229 standard; protein; 1383 AA.
 AC AAW33229;
 XX
 XX 21-OCT-2004 (revised)
 DT 30-APR-1998 (first entry)
 XX
 XX Procoagulant-active human factor VIII:C (FVIII) mutant protein.
 XX
 XX Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 FH Key
 FT Region
 FT
 FT Domain
 FT Domain
 FT Domain
 FT Disulfide-bond
 FT Domain
 FT
 FT Misc-difference
 FT
 FT Cleavage-site
 FT
 FT Domain
 FT Domain
 FT Disulfide-bond
 FT Domain
 FT Region
 FT
 FT Misc-difference
 FT
 FT Region
 FT
 FT Domain
 FT Domain

Location/Qualifiers
 1..346 factor VIIIA heavy chain
 1..329 "A1 domain"
 1..179 "plastocyanin-like domain 1"
 153..179 "probable"
 187..329 "plastocyanin-like domain 2"
 309
 /label= F309S
 /note= "wild type Phe replaced with Ser"
 372..373
 /note= "by thrombin"
 380..711
 /note= "A2 domain"
 380..554
 /note= "plastocyanin-like domain 3"
 528..554
 /note= "probable"
 564..711
 /note= "plastocyanin-like domain 4"
 711..746
 /note= "a spacer of the sequence
 SFSQSRHSTKQGFNATTIPENDIKDWF AHRTMPKIQNVSSDLMLL
 is inserted between domains A2 and A3"
 740
 /label= R740A
 /note= "wild type Arg replaced with Ala"
 741..1383
 /note= "factor VIIIA light chain"
 746..1073
 /note= "A3 domain"

FT Domain 1073..1221
 FT /note= "C1 domain"
 FT 1226..1378
 FT /note= "C2 domain"
 XX
 PN WO9740145-AL.
 XX
 PD 30-OCT-1997.
 XX
 DF 24-APR-1997; 97WO-US006563.
 XX
 PR 24-APR-1996; 96US-0016117P.
 XX 15-MAY-1996; 96US-0017785P.
 PR
 XX (UNMI) UNIV MICHIGAN.
 PA
 XX Kaufman RJ, Pipe SW, Amano K;
 FI WPI; 1997-535830/49.
 XX
 DR Modified human pro-coagulant active factor VIII - can be administered to
 FT haemophiliacs, i.e. factor VIII replacement therapy.
 PT
 XX Claim 19; Page; 57pp; English.
 PS
 XX The present sequence represents a novel pro-coagulant active factor VIII
 CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 CC Willebrand factor binding site, mutations F309S, R740A and addition of an
 CC amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
 CC along with calcium and phospholipid, acts as a cofactor for factor IXA,
 CC when it converts factor X to the activated form (factor XA). FVIII is the
 CC coagulation factor deficient in the X-chromosome-linked bleeding disorder
 CC haemophilia A. Several other mutant FVIII proteins have also been created
 CC (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
 CC recombinant secretion at higher levels than typically obtained with wild
 CC type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
 CC (AAW33222) and R562K (AAW33223) are resistant to activated protein C
 CC (APC) cleavage. The present FVIII mutant can form a more stable
 CC configuration, and have an approximate 5-fold increase in specific
 CC activity compared to purified wild type FVIII, while increasing their
 CC binding affinity to von Willebrand factor improves their stability. The
 CC FVIII proteins can be administered to haemophiliacs, i.e. FVIII
 CC replacement therapy, while the nucleic acid molecule can be used for gene
 CC therapy. note: this sequence does not appear in the specification; it was
 CC created using sequences from the given references
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to feature table key
 CC
 XX Sequence 1383 AA;
 SQ
 Query Match 100.0%; Score 1666; DB 2; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 6.7e-168; Indels 0; Gaps 0;
 Matches 313; Conservative 0; Mismatches 0;
 QY 1 LIGELHAGMSTLFLVYSNKCOTPLGMASGHIRDFQITASGOYQWAPKLARLHYSGSN 60
 DB 1052 LIGELHAGMSTLFLVYSNKCOTPLGMASGHIRDFQITASGOYQWAPKLARLHYSGSN 1111
 QY 61 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
 DB 1112 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 1171
 QY 121 GTLMVFFGNVDSSGIKHNIENPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 180
 DB 1172 GTLMVFFGNVDSSGIKHNIENPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 1231
 QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPOVNNPKWLOVDFOKTMKV 240
 DB 1232 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPOVNNPKWLOVDFOKTMKV 1291
 QY 241 TGVTTQGVKSLLTSMYKKEFLISSQDGHQWTLFPQNGKVKVFOGQNDSTFPVNSLDPP 300
 DB 1292 TGVTTQGVKSLLTSMYKKEFLISSQDGHQWTLFPQNGKVKVFOGQNDSTFPVNSLDPP 1351

Qy 301 LLTRYLRIHQPSW 313
 Db 1352 LLTRYLRIHQPSW 1364

RESULT 9
 ADQ37613
 ID ADQ37613 standard; protein; 1421 AA.
 XX
 AC ADQ37613;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human Factor VIII deletion Pro739phe mutant dB747.
 XX
 KW Human; Factor VIII; clotting factor; blood; blood disorder;
 KW haemophilia A; gene therapy; mutant; mutein; B domain.
 XX
 OS Homo sapiens.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 739
 FT Misc-difference /note= "Wild-type Pro substituted by Phe"
 FT Misc-difference 747..748
 FT Misc-difference /note= "Amino acids 748-1658 of the wild-type sequence
 have been deleted"

US2004147436-A1.
 XX
 PD 29-JUL-2004.
 XX
 XX 28-JAN-2003; 2003US-00353753.
 XX
 XX 28-JAN-2003; 2003US-00353753.
 XX
 PA (KIMH/) KIM H.
 PA (SONG/) SONG I.
 PA (CHOI/) CHOI J W.
 PA (JANG/) JANG J.
 PA (KIMY/) KIM Y.
 PA (LEE/) LEE H S.
 PA (BANG/) BANG Y.
 PA (KIMD/) KIM D.

Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
 WPI; 2004-561406/54.

New Factor VIII polypeptides that are more stable than full-length Factor VIII, useful for clotting blood or for preventing or treating blood disorders, such as Hemophilia A.

Example 4; Page; 39pp; English.

The invention relates to a Factor VIII polypeptide comprising an internal deletion of one or more amino acids between 1649 and 1688 fused to any amino acid sequence in B domain from about 741-782, with reference to full-length human Factor VIII appearing as ADQ37564. Also included are a pharmaceutical composition comprising the Factor VIII polypeptide (and a pharmaceutical carrier), a lyophilised composition comprising the Factor VIII polypeptide, clotting blood in a subject (comprising contacting a clotting amount of the Factor VIII polypeptide with the blood), treating Haemophilia A in a patient (comprising administering a clotting amount of the above Factor VIII polypeptide to a patient), an isolated nucleic acid encoding the Factor VIII polypeptide, an expression vector comprising the nucleic acid operably linked to a promoter, a host cell comprising the expression vector, making the Factor VIII polypeptide (comprising culturing the cell in conditions suitable for the vector to express the polypeptide, and isolating the polypeptide) and a purified antibody specific for the Factor VIII polypeptide cited above. The internal deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,

CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
 CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
 CC chain. The proline at 739 is replaced by another amino acid. The
 CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
 CC encompassing fusion sites between Asn amino acid at positions 745, 757 or
 CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
 CC reference to the full-length human factor VIII amino acid sequence cited
 CC above. The polypeptide comprises a formula with the following linked
 CC domains: H-S-L, where the H domain represents a polypeptide sequence
 CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
 CC the S domain represents a polypeptide spacing linker comprising up to
 CC about 60 amino acids and the L domain represents a polypeptide sequence
 CC comprising Arg-1689 through Tyr-2332 of the human factor VIII. The
 CC composition and methods are useful for clotting blood or for preventing
 CC or treating blood disorders, such as Haemophilia A. The present sequence
 CC is a human factor VIII B domain deletion mutant also carrying the
 CC Pro739phe mutation. NOTE: The present sequence is not shown in the
 CC specification but was created by the indexer using information in the
 CC specification.

XX
 SQ Sequence 1421 AA;

Query Match 100.0%; Score 1666; DB 8; Length 1421;
 Best Local Similarity 100.0%; Pred. No. 6.9e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGELHAGMSTLFLVYSNKCQTPLGWSGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
 Db 1090 LIGELHAGMSTLFLVYSNKCQTPLGWSGHIRDFQITASQYQGWAPKLARLHYSGSIN 1149

Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKQWYRGNST 120
 Db 1150 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKQWYRGNST 1209

Qy 121 GTLMVFFGNVDSSGIGKHINFPPIIARIYRLHPTHYSIRSLRMELMCDLNSCMPILGM 180
 Db 1210 GTLMVFFGNVDSSGIGKHINFPPIIARIYRLHPTHYSIRSLRMELMCDLNSCMPILGM 1269

Qy 181 ESKAISDAQITASSYFTNMPATWSPSKARLHLOGRSNARVQVNNPKEWLVDFQKTMKV 240
 Db 1270 ESKAISDAQITASSYFTNMPATWSPSKARLHLOGRSNARVQVNNPKEWLVDFQKTMKV 1329

Qy 241 TGVTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDP 300
 Db 1330 TGVTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDP 1389

Qy 301 LLTRYLRIHQPSW 313
 Db 1390 LLTRYLRIHQPSW 1402

RESULT 10
 ADQ37598
 ID ADQ37598 standard; protein; 1421 AA.
 XX
 AC ADQ37598;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human Factor VIII deletion mutant dB747.
 XX
 KW Human; Factor VIII; clotting factor; blood; blood disorder;
 KW haemophilia A; gene therapy; mutant; mutein; B domain.
 XX
 OS Homo sapiens.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 747..748
 FT Misc-difference /note= "Amino acids 748-1658 of the wild-type sequence
 have been deleted"

US2004147436-A1.

CC full-length human Factor VIII appearing as ADQ37564. Also included are a
 CC pharmaceutical composition comprising the Factor VIII polypeptide (and a
 CC pharmaceutical carrier), a lyophilised composition comprising the Factor
 CC VIII polypeptide, clotting blood in a subject (comprising contacting a
 CC clotting amount of the Factor VIII polypeptide with the blood), treating
 CC Haemophilia A in a patient (comprising administering a clotting amount of
 CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
 CC encoding the Factor VIII polypeptide, an expression vector comprising the
 CC nucleic acid operably linked to a promoter, a host cell comprising the
 CC expression vector, making the Factor VIII polypeptide (comprising
 CC culturing the cell in conditions suitable for the vector to express the
 CC polypeptide, and isolating the polypeptide) and a purified antibody
 CC specific for the Factor VIII polypeptide cited above. The internal
 CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
 CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
 CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
 CC chain. The proline at 739 is replaced by another amino acid. The
 CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
 CC encompassing fusion sites between Asn amino acid at positions 745,757 or
 CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
 CC reference to the full-length human Factor VIII amino acid sequence cited
 CC above. The polypeptide comprises a formula with the following linked
 CC domains: H-S-L, where the H domain represents a polypeptide sequence
 CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
 CC the S domain represents a polypeptide spacing linker comprising up to
 CC about 60 amino acids and the L domain represents a polypeptide sequence
 CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
 CC composition and methods are useful for clotting blood or for preventing
 CC or treating blood disorders, such as Haemophilia A. The present sequence
 CC is a human factor VIII deletion mutant where the deletion creates an N-
 CC glycosylation site. NOTE: The present sequence is not shown in the
 CC specification but was created by the indexer using information in the
 CC specification.

XX SQ Sequence 1422 AA;

Query Match 100.0%; Score 1666; DB 8; Length 1422;
 Best Local Similarity 100.0%; Pred. No. 7e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LIGHLHAGMTFLVYNSKNCQTPLGMSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 60
 DB 1091 LIGHLHAGMTFLVYNSKNCQTPLGMSGHIRDFQITASGOYQGWAPKLARLHYSGSIN 1150
 QY 61 AWSTKEPFSWIKVDLLAPMIHIGIKTQARQKFSLSYISQFLIMYSLDGKKWQYRGNST 120
 DB 1151 AWSTKEPFSWIKVDLLAPMIHIGIKTQARQKFSLSYISQFLIMYSLDGKKWQYRGNST 1210
 QY 121 GTLMVFFGNVDSGIGKHNFNPPIIARIYRLHPTHYSIRSTRMELMCGDLNCSMPGLGM 180
 DB 1211 GTLMVFFGNVDSGIGKHNFNPPIIARIYRLHPTHYSIRSTRMELMCGDLNCSMPGLGM 1270
 QY 181 ESKAISDAQITASSYFTNMFATSPSKARLHLOGESNARWQVNNPKWLDVQKTKWY 240
 DB 1271 ESKAISDAQITASSYFTNMFATSPSKARLHLOGESNARWQVNNPKWLDVQKTKWY 1330
 QY 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVQGNQDSFTPVVNSLDPP 300
 DB 1331 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVQGNQDSFTPVVNSLDPP 1390
 QY 301 LLTRYLRHPQSW 313
 DB 1391 LLTRYLRHPQSW 1403

RESULT 12
 ADQ37606
 ID ADQ37606 standard; protein; 1422 AA.
 XX AC ADQ37606;
 XX DT 21-OCT-2004 (first entry)
 XX

DE Human Factor VIII Glycosylation/deletion Pro739Phe mutant dBN(45-56).
 KW Human; Factor VIII; clotting factor; blood; blood disorder;
 XX haemophilia A; gene therapy; mutant; mutein; N-glycosylation.
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 739 /note= "Wild-type Pro substituted by Phe"
 FT Modified-site 745..747 /note= "Asn is N-glycosylated"
 FT Misc-difference 745..746 /note= "Amino acids 746-1655 of the wild-type sequence
 FT have been deleted"
 XX
 PN US2004147436-A1.
 XX
 PD 29-JUL-2004.
 XX
 XX 28-JAN-2003; 2003US-00353753.
 XX
 XX 28-JAN-2003; 2003US-00353753.
 XX
 XX (KIMH//) KIM H.
 PA (SONG//) SONG I.
 PA (CHOI//) CHOI J W.
 PA (JANG//) JANG J.
 PA (KIMY//) KIM Y.
 PA (LEE//) LEE H S.
 PA (BANG//) BANG Y.
 PA (KIMD//) KIM D.
 XX
 PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
 XX
 DR WPI; 2004-561406/54.
 XX
 XX New Factor VIII polypeptides that are more stable than full-length Factor
 PT VIII, useful for clotting blood or for preventing or treating blood
 PT disorders, such as Hemophilia A.
 XX
 PS Example 4; Page; 39pp; English.
 XX

The invention relates to a Factor VIII polypeptide comprising an internal
 deletion of one or more amino acids between 1649 and 1688 fused to any
 amino acid sequence in B domain from about 741-782, with reference to a
 full-length human factor VIII appearing as ADQ37564. Also included are a
 pharmaceutical composition comprising the Factor VIII polypeptide (and a
 pharmaceutical carrier), a lyophilised composition comprising the Factor
 VIII polypeptide, clotting blood in a subject (comprising contacting a
 clotting amount of the Factor VIII polypeptide with the blood), treating
 Haemophilia A in a patient (comprising administering a clotting amount of
 the above Factor VIII polypeptide to a patient), an isolated nucleic acid
 encoding the Factor VIII polypeptide, an expression vector comprising the
 nucleic acid operably linked to a promoter, a host cell comprising the
 expression vector, making the Factor VIII polypeptide (comprising
 culturing the cell in conditions suitable for the vector to express the
 polypeptide, and isolating the polypeptide) and a purified antibody
 specific for the Factor VIII polypeptide cited above. The internal
 deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
 CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
 CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
 CC chain. The proline at 739 is replaced by another amino acid. The
 CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
 CC encompassing fusion sites between Asn amino acid at positions 745,757 or
 CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
 CC reference to the full-length human Factor VIII amino acid sequence cited
 CC above. The polypeptide comprises a formula with the following linked
 CC domains: H-S-L, where the H domain represents a polypeptide sequence
 CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
 CC the S domain represents a polypeptide spacing linker comprising up to
 CC about 60 amino acids and the L domain represents a polypeptide sequence

CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
 CC composition and methods are useful for clotting blood or for preventing
 CC or treating blood disorders, such as haemophilia A. The present sequence
 CC is a human factor VIII deletion mutant where the deletion creates an N-
 CC glycosylation site and carries the pro739phe mutation. NOTE: The present
 CC sequence is not shown in the specification but was created by the indexer
 CC using information in the specification.
 XX
 XX Sequence 1422 AA;

Query Match 100.0%; Score 1666; DB 8; Length 1422;
 Best Local Similarity 100.0%; Pred. No. 7e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LIGELHAGMSTFLVYSNKCQTPLGWSGHIRDFOITASGQYQWAPKLARLHYSGIN 60
 DB 1091 LIGELHAGMSTFLVYSNKCQTPLGWSGHIRDFOITASGQYQWAPKLARLHYSGIN 1150
 QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFPSSLYISQFIIMYSLDGKKWQTYRGNST 120
 DB 1151 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFPSSLYISQFIIMYSLDGKKWQTYRGNST 1210
 QY 121 GTLMVFFGNVDSSGIGKNIENPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 180
 DB 1211 GTLMVFFGNVDSSGIGKNIENPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 1270
 QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 240
 DB 1271 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 1330
 QY 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300
 DB 1331 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 1390
 QY 301 LLTRYLRHPOSW 313
 DB 1391 LLTRYLRHPOSW 1403

RESULT 13
 AAP80268
 ID AAP80268 standard; protein; 1424 AA.
 XX
 XX AAP80268;
 XX
 XX 25-MAR-2003 (revised)
 DT 10-OCT-1990 (first entry)
 XX
 XX Modified factor VIII:C sequence with the R740-D1658 deletion.
 XX Modified factor VIII:C; maturation polypeptide; haemophilia;
 KW blood coagulation; RD deletion; procoagulant.
 XX
 XX Homo sapiens.
 XX
 XX WO8800831-A.
 XX
 XX 11-FEB-1988.
 XX
 XX 31-JUL-1987; 87WO-US001814.
 XX
 XX 01-AUG-1986; 86US-00893375.
 XX
 XX (BIOJ) BIOGEN NV.
 PA (PASE/) PASEK M P.
 XX
 XX Pasek MP;
 XX
 XX WPI; 1988-049866/07.
 DR N-PSDB; AAN80447.
 DR
 XX New DNA sequences encoding modified factor VIII:C - with deletion of DNA
 PT encoding maturation polypeptide, useful for high yield transformation.

XX Claim 3; Page 60-61-62-63; 97pp; English.
 XX
 XX The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part of
 CC the sequence encoding the maturation polypeptide of factor VIII:C is
 CC deleted, i.e. Gln 744 - Asp 1563. The full length Factor VIII:C cDNA has
 CC two changes with respect to the published sequence (EPO application
 CC 160457): CTG to CTA at Leu 242 and TTC to CTC change at amino acid
 CC residue 1880 (Phe to Leu). The product is produced in approx. 20 times
 CC higher yields than previous recombinant produced factor VIII:C and are
 CC more easily purified. The peptide is used for treating haemophilia A,
 CC both acute and prolonged bleeding. See also AAN80444 and AAN80446.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX

Sequence 1424 AA;
 Query Match 100.0%; Score 1666; DB 1; Length 1424;
 Best Local Similarity 100.0%; Pred. No. 7e-168;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LIGELHAGMSTFLVYSNKCQTPLGWSGHIRDFOITASGQYQWAPKLARLHYSGIN 60
 DB 1093 LIGELHAGMSTFLVYSNKCQTPLGWSGHIRDFOITASGQYQWAPKLARLHYSGIN 1152
 QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFPSSLYISQFIIMYSLDGKKWQTYRGNST 120
 DB 1153 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFPSSLYISQFIIMYSLDGKKWQTYRGNST 1212
 QY 121 GTLMVFFGNVDSSGIGKNIENPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 180
 DB 1213 GTLMVFFGNVDSSGIGKNIENPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 1272
 QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 240
 DB 1273 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 1332
 QY 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300
 DB 1333 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 1392
 QY 301 LLTRYLRHPOSW 313
 DB 1393 LLTRYLRHPOSW 1405

RESULT 14
 AAP91169
 ID AAP91169 standard; protein; 1424 AA.
 XX
 XX AAP91169;
 XX
 XX 25-MAR-2003 (revised)
 DT 26-JUN-1990 (first entry)
 XX
 XX Sequence of 740 Arg-1649 Glu human Factor VIII:C.
 DE
 XX Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
 KW haemophilia A.
 XX
 XX Homo sapiens.
 XX
 XX EP306968-A.
 XX
 XX 15-MAR-1989.
 XX
 XX 09-SEP-1988; 88EP-00114769.
 XX
 XX 10-SEP-1987; 87JP-00225147.
 PR 08-APR-1988; 88JP-00085454.
 XX
 XX (KAGA) CEMO SERO THERAPEUTIC RES INST.
 PA (TEIJ) TEIJUN LTD.
 XX

PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;
XX WPI; 1989-078467/11.
DR N-PSDB; AAN90654.
XX
XX
PT Prodn. of recombinant human Factor-VIII-C - using animal cells
PT transformed with a vector contg. the gene for Factor VIII:C and a
PT promoter.
XX
XX
PS Disclosure; Fig 1; 32pp; English.
XX
XX Arg-740 of the carboxyl terminus of the H chain is directly bonded by a
CC peptide bond to Glu-1649 of the amino terminus of L chain. A prefd.
CC expression vector used to transform animal cell so that they produce
CC human Factor VIII:Cis plasmid Ad.RE.neo. The expression vector has at
CC least one promoter upstream of AAN90654. The transformants can constantly
CC and continuously produce human Factor VIII:C in high yield on a
CC commercial scale. The human Factor VIII:C so produced is considered to
CC corresp. to the smallest species of active and intact Factor VIII:C
CC molecules in the human blood plasma. It is useful for treating
CC haemophilia A patients. (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX
SQ Sequence 1424 AA;
Query Match 100.0%; Score 1666; DB 1; Length 1424;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LIGSHLAGNSTLFLVYSNKCQTPLGNASGHIRDFQITASQYQOWAPKLARLHYSGSIN 60
Db 1093 LIGSHLAGNSTLFLVYSNKCQTPLGNASGHIRDFQITASQYQOWAPKLARLHYSGSIN 1152
Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 1153 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1212
Qy 121 GTLMVFFGNVDSSGIGKHINFPPIIARYIRLHPHYISIRSTLRMELMGCDLNSCMPLGM 180
Db 1213 GTLMVFFGNVDSSGIGKHINFPPIIARYIRLHPHYISIRSTLRMELMGCDLNSCMPLGM 1272
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1273 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1332
Qy 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 300
Db 1333 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 1392
Qy 301 LLTRYLRHPQSW 313
Db 1393 LLTRYLRHPQSW 1405
RESULT 15
AAB48842
ID AAB48842 standard; protein; 1424 AA.
XX
XX AAB48842;
XX
XX 13-MAR-2001 (first entry)
XX
XX Mutant mature human factor VIII, SEQ ID NO:5.
XX
XX Factor VIII; human; B domain; LRP-mediated plasma clearance;
KW receptor-dependent clearance; LRP-mediated plasma clearance; half-life;
XX haemophilia; mutant; muten.
XX
XX Homo sapiens.
XX
XX WO200071714-A2.
XX
XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US014111.
XX
XX 24-MAY-1999; 99US-0135847P.
XX
XX (AMNA-) AMERICAN NAT RED CROSS.
XX
XX Saenko EL, Strickland DK;
XX
XX WPI; 2001-025163/03.
XX
XX Factor VIII mutants having increased half-life useful for treating
PT hemophilia, comprise one or more amino acid substitutions in the A2
PT and/or C2 domain of factor VIII.
XX
XX Claim 9; Fig 2A-B; 121pp; English.
XX
XX The invention relates to human factor VIII mutants comprising an amino
CC acid substitution at one or more positions in the A2 domain and/or an
CC amino acid substitution at one or more positions in the C2 domain. The
CC invention also encompasses a factor VIII mutant which lacks a B domain
CC (AAB48842). The factor VIII mutants have an increased half-life in the
CC bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants
CC have reduced receptor-independent clearance. The invention also relates
CC to a method of using RAP (receptor associated protein), a protein which
CC inhibits LRP (low density lipoprotein related protein)-mediated ligand
CC internalisation, to increase the half-life of factor VIII. The mutant
CC factor VIII proteins, and nucleotides encoding them, are useful for
CC treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
CC nucleic acids encoding them may also be used in the treatment of
CC haemophilia, in combination with a mutant factor VIII protein or DNA of
CC the invention. The invention provides means of increasing the half-life
CC of factor VIII by reducing its clearance from plasma. The present
CC sequence represents a mutant mature human factor VIII which lacks a B
XX domain
XX
SQ Sequence 1424 AA;
Query Match 100.0%; Score 1666; DB 4; Length 1424;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LIGSHLAGNSTLFLVYSNKCQTPLGNASGHIRDFQITASQYQOWAPKLARLHYSGSIN 60
Db 1093 LIGSHLAGNSTLFLVYSNKCQTPLGNASGHIRDFQITASQYQOWAPKLARLHYSGSIN 1152
Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 1153 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1212
Qy 121 GTLMVFFGNVDSSGIGKHINFPPIIARYIRLHPHYISIRSTLRMELMGCDLNSCMPLGM 180
Db 1213 GTLMVFFGNVDSSGIGKHINFPPIIARYIRLHPHYISIRSTLRMELMGCDLNSCMPLGM 1272
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1273 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1332
Qy 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 300
Db 1333 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 1392
Qy 301 LLTRYLRHPQSW 313
Db 1393 LLTRYLRHPQSW 1405
Search completed: April 20, 2005, 17:19:11
Job time : 99.6188 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:14:28 ; Search time 30.2153 Seconds
(without alignments)
773.288 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGSHLAGMSTFLVYSNK.....VNSLDPLLTLRLHPOSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1666	100.0	1438	3	US-09-209-916-1
2	1666	100.0	1457	4	US-09-001-039B-47
3	1666	100.0	1471	1	US-08-683-839B-3
4	1666	100.0	1661	2	US-08-882-083-2
5	1666	100.0	1661	2	US-08-558-107-2
6	1666	100.0	1661	3	US-09-243-539-2
7	1666	100.0	2332	1	US-07-864-004B-4
8	1666	100.0	2332	1	US-08-251-937A-4
9	1666	100.0	2332	1	US-08-212-133A-2
10	1666	100.0	2332	1	US-08-474-503-2
11	1666	100.0	2332	2	US-08-670-707A-2
12	1666	100.0	2332	3	US-09-037-601-2
13	1666	100.0	2332	3	US-09-324-867-3
14	1666	100.0	2332	3	US-09-315-179-2
15	1666	100.0	2332	4	US-09-523-656-2
16	1666	100.0	2332	4	US-09-957-641A-2
17	1666	100.0	2332	5	PCT-US93-03275-4
18	1666	100.0	2332	5	PCT-US94-13200-2
19	1666	100.0	2351	1	US-08-121-202-2
20	1666	100.0	2351	1	US-08-366-851A-2
21	1666	100.0	2351	4	US-10-133-907-4
22	1666	100.0	2351	4	US-09-001-039B-45
23	1663	99.8	2332	1	US-08-276-594A-2
24	1660	99.6	2351	6	5171844-2
25	1660	99.6	2351	6	5171844-2
26	1657	99.5	2351	6	5422260-1
27	1657	99.5	2351	6	5422260-1

28	1522	91.4	2343	3	US-09-324-867-2	Sequence 2, Appli
29	1495	89.7	2319	1	US-08-212-133A-8	Sequence 8, Appli
30	1495	89.7	2319	1	US-08-474-503-6	Sequence 6, Appli
31	1495	89.7	2319	2	US-08-670-707A-6	Sequence 6, Appli
32	1495	89.7	2319	3	US-09-037-601-6	Sequence 6, Appli
33	1495	89.7	2319	3	US-09-315-179-6	Sequence 6, Appli
34	1495	89.7	2319	4	US-09-523-656-28	Sequence 28, Appli
35	1495	89.7	2319	5	PCT-US94-13200-6	Sequence 6, Appli
36	1480.5	88.9	2304	3	US-09-324-867-4	Sequence 4, Appli
37	1467	88.1	1443	2	US-08-670-707A-39	Sequence 39, Appli
38	1467	88.1	1443	3	US-09-037-601-39	Sequence 39, Appli
39	1467	88.1	1443	3	US-09-315-179-39	Sequence 39, Appli
40	1467	88.1	1467	4	US-09-523-656-38	Sequence 38, Appli
41	1467	88.1	2133	2	US-08-670-707A-37	Sequence 37, Appli
42	1467	88.1	2133	3	US-09-037-601-37	Sequence 37, Appli
43	1467	88.1	2133	4	US-09-315-179-37	Sequence 37, Appli
44	1467	88.1	2133	4	US-09-523-656-30	Sequence 30, Appli
45	1454	87.3	2115	3	US-09-324-867-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

Query Match	100.0%	Score	1666;	DB	3;	Length	1438;
Best Local Similarity	100.0%	Pred. No.	9.8e-168;	Mismatches	0;	Indels	0;
Matches	313;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	LIGSHLAGMSTFLVYSNKQTPLGMSGHIRDFQITASQYQGWAPKARLHYSGSIN	60				
Db	1107	LIGSHLAGMSTFLVYSNKQTPLGMSGHIRDFQITASQYQGWAPKARLHYSGSIN	1166				
QY	61	AWSTKEPSPWIKVDLLAPMIITHGIKTQAGRKFSLSYISQFIIMYSLDCKWKQTYRGNST	120				
Db	1167	AWSTKEPSPWIKVDLLAPMIITHGIKTQAGRKFSLSYISQFIIMYSLDCKWKQTYRGNST	1226				
QY	121	GLMVFFGNVDSSGKINIFNPPIIARYIRLHPHYSTIRSLRMELMCDLNSCSMPLGM	180				
Db	1227	GLMVFFGNVDSSGKINIFNPPIIARYIRLHPHYSTIRSLRMELMCDLNSCSMPLGM	1286				
QY	181	ESKAISDAQITASSYFTNMFTATSPSKARLHQGRSNARWQVNNPKWLQVDFQTKWKV	240				
Db	1287	ESKAISDAQITASSYFTNMFTATSPSKARLHQGRSNARWQVNNPKWLQVDFQTKWKV	1346				
QY	241	TGVTQGVKSLTSMYKKEFLISSSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDDPP	300				
Db	1347	TGVTQGVKSLTSMYKKEFLISSSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDDPP	1406				
QY	301	LLTRYLRHPOSW	313				

Db 1407 LLTRYLRIHPQSW 1419

RESULT 2

US-09-001-039B-47

; Sequence 47, Application US/09001039B

; Patent No. 6818439

; GENERAL INFORMATION:

; APPLICANT: Jolly, Douglas J.

; APPLICANT: Chang, Stephen M.W.

; APPLICANT: Respass, James G.

; APPLICANT: DePolo, Nicholas J.

; APPLICANT: Hsu, David Chi-Tang

; APPLICANT: Ibanez, Carlos E.

; APPLICANT: Greengard, Judith

; APPLICANT: Lee, Will

; TITLE OF INVENTION: METHODS FOR ADMINISTRATION OF

; TITLE OF INVENTION: RECOMBINANTGENE DELIVERY VEHICLES FOR TREATMENT

; TITLE OF INVENTION: OF HEMOPHILIA AND OTHER DISORDERS

; NUMBER OF SEQUENCES: 84

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed Intellectual Property Law Group

; STREET: 701 Fifth Avenue, Suite 6300

; CITY: Seattle

; STATE: Washington

; COUNTRY: U.S.A.

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/001,039B

; FILING DATE: 13-JAN-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: McWaters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 1155.005 / 930049.441C4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1457 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-001-039B-47

Query Match 100.0%; Score 1666; DB 4; Length 1457;

Best Local Similarity 100.0%; Pred. No. 1e-167;

Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFOITASGYGQWAPKRLARLHYSGSIN 60

Db 1126 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFOITASGYGQWAPKRLARLHYSGSIN 1185

QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKPFSSLYISQFIIMYSLDGKKWQTYRGNST 120

Db 1186 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKPFSSLYISQFIIMYSLDGKKWQTYRGNST 1245

QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 180

Db 1246 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 1305

QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARVPQVNNPKWLOVDFOKTMKV 240

Db 1306 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARVPQVNNPKWLOVDFOKTMKV 1365

QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDPP 300

Db 1366 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDPP 1425

QY 301 LLTRYLRIHPQSW 313

Db 1426 LLTRYLRIHPQSW 1438

RESULT 3

US-08-683-839B-3

; Sequence 3, Application US/08683839B

; Patent No. 5744326

; GENERAL INFORMATION:

; APPLICANT: Ill, Charles . R. et al.

; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional

; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of

; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street, suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/683,839B

; FILING DATE: 11-MARCH-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Remillard, Jane E.

; REGISTRATION NUMBER: 38,872

; REFERENCE/DOCKET NUMBER: TTI-138

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1471 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-683-839B-3

Query Match 100.0%; Score 1666; DB 1; Length 1471;

Best Local Similarity 100.0%; Pred. No. 1e-167;

Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFOITASGYGQWAPKRLARLHYSGSIN 60

Db 1140 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFOITASGYGQWAPKRLARLHYSGSIN 1199

QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKPFSSLYISQFIIMYSLDGKKWQTYRGNST 120

Db 1200 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKPFSSLYISQFIIMYSLDGKKWQTYRGNST 1259

QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 180

Db 1260 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPHTHYSIRSLRMLMGCCLNSCSMPLGM 1319

QY 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARVPQVNNPKWLOVDFOKTMKV 240

Db 1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARVPQVNNPKWLOVDFOKTMKV 1379

QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTFPVNSLDPP 300

Db 1380 TGVTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 1439
 Qy 301 LLTRYLRIHQSW 313
 Db 1440 LLTRYLRIHQSW 1452

RESULT 4
 US-08-882-083-2
 ; Sequence 2, Application US/08882083
 ; Patent No. 5869292
 ; GENERAL INFORMATION:
 ; APPLICANT: VOORBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/882,083
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/558,107
 ; FILING DATE: 13-NOV-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ISACSON, John P.
 ; REGISTRATION NUMBER: 33,715
 ; REFERENCE/DOCKET NUMBER: 30472/212
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-882-083-2

Query Match 100.0%; Score 1666; DB 2; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 1.2e-167;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
 Db 1330 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSIN 1389
 Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
 Db 1390 AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 1449
 Qy 121 GTLMVFNGVNDSSGKINIFNPPIIARYIRLHPHYISIRSLRMLMGCDLNSCMLPLGM 180
 Db 1450 GTLMVFNGVNDSSGKINIFNPPIIARYIRLHPHYISIRSLRMLMGCDLNSCMLPLGM 1509
 Qy 181 ESKAISDAQITASSYFTNMFTATSPSKARLHQRSNARVPQVNNPKWQLQVDFQKTMKV 240
 Db 1510 ESKAISDAQITASSYFTNMFTATSPSKARLHQRSNARVPQVNNPKWQLQVDFQKTMKV 1569
 Qy 241 TGVTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300

Db 1570 TGVTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 1629
 Qy 301 LLTRYLRIHQSW 313
 Db 1630 LLTRYLRIHQSW 1642

RESULT 5
 US-08-558-107-2
 ; Sequence 2, Application US/08558107
 ; Patent No. 5910481
 ; GENERAL INFORMATION:
 ; APPLICANT: VOORBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/558,107
 ; FILING DATE: 13-NOV-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ISACSON, John P.
 ; REGISTRATION NUMBER: 33,715
 ; REFERENCE/DOCKET NUMBER: 30472/212
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-558-107-2

Query Match 100.0%; Score 1666; DB 2; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 1.2e-167;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
 Db 1330 LIGELHAGMSTLFLVYSNKCOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSIN 1389
 Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
 Db 1390 AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 1449
 Qy 121 GTLMVFNGVNDSSGKINIFNPPIIARYIRLHPHYISIRSLRMLMGCDLNSCMLPLGM 180
 Db 1450 GTLMVFNGVNDSSGKINIFNPPIIARYIRLHPHYISIRSLRMLMGCDLNSCMLPLGM 1509
 Qy 181 ESKAISDAQITASSYFTNMFTATSPSKARLHQRSNARVPQVNNPKWQLQVDFQKTMKV 240
 Db 1510 ESKAISDAQITASSYFTNMFTATSPSKARLHQRSNARVPQVNNPKWQLQVDFQKTMKV 1569
 Qy 241 TGVTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300
 Db 1570 TGVTTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 1629
 Qy 301 LLTRYLRIHQSW 313

Db 1630 LLTRYLRHPQSW 1642

RESULT 6
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-243-539-2

Query Match 100.0%; Score 1666; DB 3; Length 1661;
Best Local Similarity 100.0%; Pred. No. 1.2e-167; Indels 0; Gaps 0;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LIGEHLAGMSTFLVYSNKCTPLGASGHIRDFQITASGOYGOWAPKRLARLHYSGSIN 60
Db 1330 LIGEHLAGMSTFLVYSNKCTPLGASGHIRDFQITASGOYGOWAPKRLARLHYSGSIN 1389
Qy 61 AWSTKEPFSWKVDLLAPMIHGIKTQAGRKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 1390 AWSTKEPFSWKVDLLAPMIHGIKTQAGRKFSLSYISQFIIMYSLDGKKWQTYRGNST 1449
Qy 121 GTLMVFFGNVDSGGIKHNFNPPIIARYIRLHPHYSIRSTLRMELMGCDLNSCSMPLGM 180
Db 1450 GTLMVFFGNVDSGGIKHNFNPPIIARYIRLHPHYSIRSTLRMELMGCDLNSCSMPLGM 1509
Qy 181 ESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNAWRPOVNNPKEWLQVDFOKTMKV 240
Db 1510 ESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNAWRPOVNNPKEWLQVDFOKTMKV 1569
Qy 241 TGVTTQGVKSLTSMVYKEFLISSODGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 1570 TGVTTQGVKSLTSMVYKEFLISSODGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 1629
Qy 301 LLTRYLRHPQSW 313
Db 1630 LLTRYLRHPQSW 1642

RESULT 7
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
; US-07-864-004B-4
Query Match 100.0%; Score 1666; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167; Indels 0; Gaps 0;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LIGEHLAGMSTFLVYSNKCTPLGASGHIRDFQITASGOYGOWAPKRLARLHYSGSIN 60
Db 2001 LIGEHLAGMSTFLVYSNKCTPLGASGHIRDFQITASGOYGOWAPKRLARLHYSGSIN 2060
Qy 61 AWSTKEPFSWKVDLLAPMIHGIKTQAGRKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWKVDLLAPMIHGIKTQAGRKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120
Qy 121 GTLMVFFGNVDSGGIKHNFNPPIIARYIRLHPHYSIRSTLRMELMGCDLNSCSMPLGM 180
Db 2121 GTLMVFFGNVDSGGIKHNFNPPIIARYIRLHPHYSIRSTLRMELMGCDLNSCSMPLGM 2180
Qy 181 ESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNAWRPOVNNPKEWLQVDFOKTMKV 240
Db 2181 ESKAISDAQITASSYFTNNFATWSPSKARLHLOGRSNAWRPOVNNPKEWLQVDFOKTMKV 2240
Qy 241 TGVTTQGVKSLTSMVYKEFLISSODGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMVYKEFLISSODGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 2300

```

Qy 301 LLTRYLRHPSW 313
Db 2301 LLTRYLRHPSW 2313

RESULT 8
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
US-08-251-937A-4

Query Match 100.0%; Score 1666; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTFLVYSNKKQTPLGMSAGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTFLVYSNKKQTPLGMSAGHIRDFQITASQYQGWAPKLARLHYSGSIN 2060
Qy 61 AWSKPEPSWIKVDLLAPMIHGIKTQCAROKFSSLYISQFIIMYSLDGKWKQTVRGNST 120
Db 2061 AWSKPEPSWIKVDLLAPMIHGIKTQCAROKFSSLYISQFIIMYSLDGKWKQTVRGNST 2120
Qy 121 GTLMVFFGNVDSSGIGKNIIFNPPIIARIYIRLHPHYISIRSLRMLMCDLNSCMPILGM 180
Db 2121 GTLMVFFGNVDSSGIGKNIIFNPPIIARIYIRLHPHYISIRSLRMLMCDLNSCMPILGM 2180
Qy 181 ESKAISDAQITASSYFTNMFATWSFKARLHQLGRSNARPOVNNPKEWLQVDFQTKMKV 240
Db 181 ESKAISDAQITASSYFTNMFATWSFKARLHQLGRSNARPOVNNPKEWLQVDFQTKMKV 240

Qy 2181 ESKAISDAQITASSYFTNMFATWSFKARLHQLGRSNARPOVNNPKEWLQVDFQTKMKV 2240
Qy 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDQSFPTPVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDQSFPTPVNSLDPP 2300

Qy 301 LLTRYLRHPSW 313
Db 2301 LLTRYLRHPSW 2313

RESULT 9
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabat, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
US-08-212-133A-2

Query Match 100.0%; Score 1666; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTFLVYSNKKQTPLGMSAGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTFLVYSNKKQTPLGMSAGHIRDFQITASQYQGWAPKLARLHYSGSIN 2060
Qy 61 AWSKPEPSWIKVDLLAPMIHGIKTQCAROKFSSLYISQFIIMYSLDGKWKQTVRGNST 120
Db 2061 AWSKPEPSWIKVDLLAPMIHGIKTQCAROKFSSLYISQFIIMYSLDGKWKQTVRGNST 2120

```

QY 121 GTLMVFFGVNDSGKHNIFNPPIIARYIRLHPHYSIRSTRMLMELMGCDLNSCSMPLGM 180
 Db 2121 GTLMVFFGVNDSGKHNIFNPPIIARYIRLHPHYSIRSTRMLMELMGCDLNSCSMPLGM 2180
 QY 181 ESKAISDAQITASSYFTNMFTATSPSKARLHLOGRSNARPOVNNPKWLQVDFQKTMKV 240
 Db 2181 ESKAISDAQITASSYFTNMFTATSPSKARLHLOGRSNARPOVNNPKWLQVDFQKTMKV 2240
 QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDQSFPPVNSLDPP 300
 Db 2241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDQSFPPVNSLDPP 2300
 QY 301 LLTRYLRHQPQSW 313
 Db 2301 LLTRYLRHQPQSW 2313
 RESULT 10
 US-08-474-503-2
 ; Sequence 2, Application US/08474503
 ; Patent No. 5744446
 ; GENERAL INFORMATION:
 ; APPLICANT: Emory University
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kilpatrick & Cody
 ; STREET: 1100 Peachtree Street, Suite 2800
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: US
 ; ZIP: 30309
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,503
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pratt, John S.
 ; REGISTRATION NUMBER: 29,476
 ; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404-815-6500
 ; TELEFAX: 404-815-6555
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2332 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapien
 ; TISSUE TYPE: Liver cdna sequence
 ; US-08-474-503-2
 Query Match 100.0%; Score 1666; DB 1; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 2.1e-167;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LIGELHAGMSTLFLVYSNKCOTPIGMAHGIRDFQTASQYQWAPKRLHYSGSIN 60
 Db 2001 LIGELHAGMSTLFLVYSNKCOTPIGMAHGIRDFQTASQYQWAPKRLHYSGSIN 2060
 QY 61 AWSTKPFPSWIKVDLLAPMIHGIKTQCAROKFSSLIYSQFIIMYSLDGGKKWQTVRGNST 120

Db 2061 AWSTKPFPSWIKVDLLAPMIHGIKTQCAROKFSSLIYSQFIIMYSLDGGKKWQTVRGNST 2120
 QY 121 GTLMVFFGVNDSGKHNIFNPPIIARYIRLHPHYSIRSTRMLMELMGCDLNSCSMPLGM 180
 Db 2121 GTLMVFFGVNDSGKHNIFNPPIIARYIRLHPHYSIRSTRMLMELMGCDLNSCSMPLGM 2180
 QY 181 ESKAISDAQITASSYFTNMFTATSPSKARLHLOGRSNARPOVNNPKWLQVDFQKTMKV 240
 Db 2181 ESKAISDAQITASSYFTNMFTATSPSKARLHLOGRSNARPOVNNPKWLQVDFQKTMKV 2240
 QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDQSFPPVNSLDPP 300
 Db 2241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNDQSFPPVNSLDPP 2300
 QY 301 LLTRYLRHQPQSW 313
 Db 2301 LLTRYLRHQPQSW 2313
 RESULT 11
 US-08-670-707A-2
 ; Sequence 2, Application US/08670707A
 ; Patent No. 5859204
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/670,707A
 ; FILING DATE: 26-JUN-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Greenlee, Winner and Sullivan, P.C.
 ; REGISTRATION NUMBER: 27,894
 ; REFERENCE/DOCKET NUMBER: 75-95F
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/439-8080
 ; TELEFAX: 303/439-8089
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2332 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; TISSUE TYPE: Liver
 ; US-08-670-707A-2

Query Match 100.0%; Score 1666; DB 2; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYVSNKQCTPLGMSGHIRDFQITASGOYQOWAPKARLHYSGSIN 60
DB 2001 LIGEHLAGMSTFLVYVSNKQCTPLGMSGHIRDFQITASGOYQOWAPKARLHYSGSIN 2060

QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQAGARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
DB 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQAGARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARIYIRLPHYTHYSIRSTRMLMCGDLNCSMPLGM 180
DB 2121 GTLMVFFGNVDSSGKHNIFNPPIIARIYIRLPHYTHYSIRSTRMLMCGDLNCSMPLGM 2180

QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHQGRSNARPOVNNPKWLOVDFOKTKV 240
DB 2181 ESKAISDAQITASSYFTNMFTWSPSKARLHQGRSNARPOVNNPKWLOVDFOKTKV 2240

QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300
DB 2241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 2300

QY 301 LLTRYLRHPOSW 313
DB 2301 LLTRYLRHPOSW 2313

RESULT 12
US-09-37-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-09-037-601-2

Query Match 100.0%; Score 1666; DB 3; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYVSNKQCTPLGMSGHIRDFQITASGOYQOWAPKARLHYSGSIN 60
DB 2001 LIGEHLAGMSTFLVYVSNKQCTPLGMSGHIRDFQITASGOYQOWAPKARLHYSGSIN 2060

QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQAGARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
DB 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQAGARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARIYIRLPHYTHYSIRSTRMLMCGDLNCSMPLGM 180
DB 2121 GTLMVFFGNVDSSGKHNIFNPPIIARIYIRLPHYTHYSIRSTRMLMCGDLNCSMPLGM 2180

QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHQGRSNARPOVNNPKWLOVDFOKTKV 240
DB 2181 ESKAISDAQITASSYFTNMFTWSPSKARLHQGRSNARPOVNNPKWLOVDFOKTKV 2240

QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 300
DB 2241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVNSLDPP 2300

QY 301 LLTRYLRHPOSW 313
DB 2301 LLTRYLRHPOSW 2313

RESULT 13
US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
; APPLICANT: Lillcrap, David
; APPLICANT: Cameron, Cherie
; APPLICANT: No. 6251632ley, Colleen
; APPLICANT: Horrocks, L. Suzanne Hoyle
; APPLICANT: Hough, Christine
; TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
; FILE REFERENCE: 1669 001002/JAG/BJD
; CURRENT APPLICATION NUMBER: US/09/324,867A
; CURRENT FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: 09/035,141
; EARLIER FILING DATE: 1998-03-059
; EARLIER APPLICATION NUMBER: 60/039,953
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-324-867-3

Query Match 100.0%; Score 1666; DB 3; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYVSNKQCTPLGMSGHIRDFQITASGOYQOWAPKARLHYSGSIN 60

```
Db 2001 LIGEHLAGMSTFLVYSNKCQTPLGASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2060
Qy 61 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120
Qy 121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 180
Db 2121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 2240
Qy 241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 2300
Qy 301 LLTRYLRIRHPQSW 313
Db 2301 LLTRYLRIRHPQSW 2313

RESULT 14
US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; EARLIER FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200
; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-179-2

Query Match 100.0%; Score 1666; DB 3; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGEHLAGMSTFLVYSNKCQTPLGASGHIRDFOITASQYQGWAPKLARLHYSGSIN 60
Db 2001 LIGEHLAGMSTFLVYSNKCQTPLGASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2060
Qy 61 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120
Qy 121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 180
Db 2121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 2240

Search completed: April 20, 2005, 17:22:32
Job time : 31.2153 secs
```

```
Qy 241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 2300
Qy 301 LLTRYLRIRHPQSW 313
Db 2301 LLTRYLRIRHPQSW 2313

RESULT 15
US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; EARLIER FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

Query Match 100.0%; Score 1666; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGEHLAGMSTFLVYSNKCQTPLGASGHIRDFOITASQYQGWAPKLARLHYSGSIN 60
Db 2001 LIGEHLAGMSTFLVYSNKCQTPLGASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2060
Qy 61 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFWSIKVDLLAPMI IHGKTQAROKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120
Qy 121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 180
Db 2121 GTLMVFFGNVDSSGIGKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLQVDFQKTMKV 2240
Qy 241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 2300
Qy 301 LLTRYLRIRHPQSW 313
Db 2301 LLTRYLRIRHPQSW 2313

Search completed: April 20, 2005, 17:22:32
Job time : 31.2153 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:21:09 ; Search time 67.4035 Seconds
(without alignments)
1543.422 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLAGMSTFLVYSNK.....VNSLDPLRLRYLRHPSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	1438	13	US-10-006-091-1
2	1666	100.0	1438	13	US-10-047-257-1
3	1666	100.0	1438	14	US-10-225-900-1
4	1666	100.0	1459	15	US-10-239-498A-13
5	1666	100.0	1471	13	US-10-095-718-2
6	1666	100.0	1471	15	US-10-681-970-2
7	1666	100.0	2332	9	US-09-957-641-2
8	1666	100.0	2332	14	US-10-187-319-2
9	1666	100.0	2332	15	US-10-131-510A-2
10	1666	100.0	2332	15	US-10-445-235-2
11	1666	100.0	2332	15	US-10-360-101-229
12	1666	100.0	2332	15	US-10-239-498A-2
13	1666	100.0	2332	16	US-10-466-998A-1

14	1666	100.0	2332	16	US-10-721-997A-34	Sequence 34, Appl
15	1666	100.0	2351	14	US-10-132-829-4	Sequence 4, Appl
16	1666	100.0	2351	14	US-10-172-712-27	Sequence 27, Appl
17	1666	100.0	2351	14	US-10-133-907-4	Sequence 4, Appl
18	1666	100.0	2351	15	US-10-411-037-30	Sequence 30, Appl
19	1666	100.0	2351	15	US-10-411-026-30	Sequence 30, Appl
20	1666	100.0	2351	15	US-10-410-962-30	Sequence 30, Appl
21	1666	100.0	2351	15	US-10-411-049-30	Sequence 30, Appl
22	1666	100.0	2351	16	US-10-410-930-30	Sequence 30, Appl
23	1666	100.0	2351	16	US-10-410-930-30	Sequence 30, Appl
24	1666	100.0	2351	16	US-10-411-012-30	Sequence 30, Appl
25	1666	100.0	2351	16	US-10-287-994-30	Sequence 30, Appl
26	1666	100.0	2351	16	US-10-410-913-30	Sequence 30, Appl
27	1666	100.0	2351	17	US-10-410-980-30	Sequence 30, Appl
28	1666	99.6	1459	15	US-10-239-498A-15	Sequence 15, Appl
29	1657	99.5	1459	15	US-10-239-498A-4	Sequence 4, Appl
30	1522	91.4	1431	13	US-10-095-718-4	Sequence 4, Appl
31	1522	91.4	1431	15	US-10-681-970-4	Sequence 4, Appl
32	1495	89.7	2319	14	US-10-187-319-6	GENERAL INFORMA
33	1495	89.7	2319	14	US-10-131-510A-6	Sequence 6, Appl
34	1484	89.1	2351	17	US-10-741-600-1034	Sequence 1034, Ap
35	1467	88.1	1443	14	US-10-187-319-39	Sequence 39, Appl
36	1467	88.1	1443	14	US-10-131-510A-39	Sequence 39, Appl
37	1467	88.1	2114	16	US-10-721-997A-33	Sequence 33, Appl
38	1467	88.1	2133	14	US-10-187-319-37	Sequence 37, Appl
39	1467	88.1	2133	14	US-10-131-510A-37	Sequence 37, Appl
40	1055	63.3	218	15	US-10-038-252-5	Sequence 5, Appl
41	1006	60.4	216	17	US-10-741-600-1033	Sequence 1033, Ap
42	748	44.9	160	15	US-10-298-796-9	Sequence 9, Appl
43	745.5	44.7	2196	15	US-10-360-101-259	Sequence 259, App
44	745.5	44.7	2224	14	US-10-115-563-14	Sequence 14, Appl
45	745.5	44.7	2224	14	US-10-172-712-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match	100.0%	Score	1666	DB	13	Length	1438
Best Local Similarity	100.0%	Pred. No.	7.1e-162				
Matches	313	Conservative	0	Mismatches	0	Indels	0
Qy	1	LIGEHLAGMSTFLVYSNKCPPLGMSGHIRDFQTASQYCGWAPKLARLYSGSIN	60				
Db	1107	LIGEHLAGMSTFLVYSNKCPPLGMSGHIRDFQTASQYCGWAPKLARLYSGSIN	1166				
Qy	61	AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKQWYRGNST	120				
Db	1167	AWSTKEPFSWIKVDLLAPMIHGIKTQCARQKFSLSYISQFIIMYSLDGKKQWYRGNST	1226				

QY 121 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 180
DB 1227 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 1286
QY 181 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 1287 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 1346
QY 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
DB 1347 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 1406
QY 301 LLTRYLRHPQSW 313
DB 1407 LLTRYLRHPQSW 1419

RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match 100.0%; Score 1666; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 7.1e-162;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIGEHLAGMSTFLVYSNKCOTPLGNWASGHIRDFOITASGOYGOWAPKLARLHYSGIN 60
DB 1107 LIGEHLAGMSTFLVYSNKCOTPLGNWASGHIRDFOITASGOYGOWAPKLARLHYSGIN 1166
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSYISQFIIMYSLDGKKWQTYRGNST 120
DB 1167 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSYISQFIIMYSLDGKKWQTYRGNST 1226
QY 121 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 180
DB 1227 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 1286
QY 181 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 1287 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 1346
QY 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
DB 1347 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 1406
QY 301 LLTRYLRHPQSW 313
DB 1407 LLTRYLRHPQSW 1419

RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900

; Publication No. US20030077752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

Query Match 100.0%; Score 1666; DB 14; Length 1438;
Best Local Similarity 100.0%; Pred. No. 7.1e-162;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIGEHLAGMSTFLVYSNKCOTPLGNWASGHIRDFOITASGOYGOWAPKLARLHYSGIN 60
DB 1107 LIGEHLAGMSTFLVYSNKCOTPLGNWASGHIRDFOITASGOYGOWAPKLARLHYSGIN 1166
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSYISQFIIMYSLDGKKWQTYRGNST 120
DB 1167 AWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSYISQFIIMYSLDGKKWQTYRGNST 1226
QY 121 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 180
DB 1227 GTLMVFFGNVDSGKHNIFNPPIIARYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 1286
QY 181 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 1287 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 1346
QY 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
DB 1347 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 1406
QY 301 LLTRYLRHPQSW 313
DB 1407 LLTRYLRHPQSW 1419

RESULT 4
US-10-239-498A-13
; Sequence 13, Application US/10239498A
; Publication No. US2004002333A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Charlotte
; APPLICANT: Horster, Andrea
; APPLICANT: Schroder, Carola
; APPLICANT: Lehner, Michael
; TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
; FILE REFERENCE: 80977.0001
; CURRENT APPLICATION NUMBER: US/10/239,498A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/EP01/03220
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1459
; TYPE: PRT


```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTGF8-2hyg-s
US-10-239-498A-13

Query Match      100.0%; Score 1666; DB 15; Length 1459;
Best Local Similarity 100.0%; Pred. No. 7.2e-162;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
Db 1128 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 1187
Qy 61 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 1188 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1247
Qy 121 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 180
Db 1248 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 1307
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1308 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1367
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 300
Db 1368 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 1427
Qy 301 LLTRYLRHQPQSW 313
Db 1428 LLTRYLRHQPQSW 1440

RESULT 5
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Methods of Using the Same
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/095,718
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

Query Match      100.0%; Score 1666; DB 13; Length 1471;
Best Local Similarity 100.0%; Pred. No. 7.3e-162;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
Db 1140 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 1199
Qy 61 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 1188 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1247
Qy 121 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 180
Db 1248 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 1307
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1308 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1367
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 300
Db 1428 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 1471
Qy 301 LLTRYLRHQPQSW 313
Db 1428 LLTRYLRHQPQSW 1440
```

```
Db 1200 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1259
Qy 121 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 180
Db 1260 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 1319
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1379
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 300
Db 1380 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 1439
Qy 301 LLTRYLRHQPQSW 313
Db 1440 LLTRYLRHQPQSW 1452

RESULT 6
US-10-681-970-2
; Sequence 2, Application US/10681970
; Publication No. US20040062752A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Munson, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/681,970
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US/09/689,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/158,780
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: PRT
; ORGANISM: Homo sapiens B-domain deleted factor VIII
; FEATURE:
; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-681-970-2

Query Match      100.0%; Score 1666; DB 15; Length 1471;
Best Local Similarity 100.0%; Pred. No. 7.3e-162;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 60
Db 1140 LIGELHAGMSTLFLVYSNKKQTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSIN 1199
Qy 61 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 1200 AWSTKEPFSWKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQYRGNST 1259
Qy 121 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 180
Db 1260 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHYSIRSTRMELMCGDLNSCSMPGLM 1319
Qy 181 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
Db 1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKEWLQVDFQKTMKV 1379
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 300
Db 1380 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGQNDQSFPTPVNSLDPP 1439
```

```
QY 301 LLTRYLRHQPQSW 313
Db 1440 LLTRYLRHQPQSW 1452

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/187,319
  FILING DATE: 27-Aug-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 09/523,656
  FILING DATE: 2000-03-10
  APPLICATION NUMBER: US 09/037,601
  FILING DATE: 1998-03-10
  APPLICATION NUMBER: WO PCT/US97/11155
  FILING DATE: 1997-06-26
  APPLICATION NUMBER: US 08/670,707
  FILING DATE: 1996-06-26
ATTORNEY/AGENT INFORMATION:
  NAME: Greenlee, Lorraine L.
  REGISTRATION NUMBER: 27,894
  REFERENCE/DOCKET NUMBER: 75-95K
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 303/499-8080
    TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2332 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: unknown
  MOLECULE TYPE: protein
  HYPOTHEICAL: YES
  ANTI-SENSE: NO
  FRAGMENT TYPE: N-terminal
  ORIGINAL SOURCE:
    ORGANISM: Homo sapiens
    TISSUE TYPE: Liver
  SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

Query Match 100.0%; Score 1666; DB 9; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYSNKCQTPLGMSGHIRDFQITASGOYGQWAPKLARLHYSGSIN 60
Db 2001 LIGELHAGMSTFLVYSNKCQTPLGMSGHIRDFQITASGOYGQWAPKLARLHYSGSIN 2060

QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 2061 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 2120

QY 121 GTLMVFFGNVDSSGKKNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFFGNVDSSGKKNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 2180

QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNARPOVNNPKEWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNARPOVNNPKEWLQVDFQKTMKV 2240

QY 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGQNDSPFPVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGQNDSPFPVNSLDPP 2300

QY 301 LLTRYLRHQPQSW 313
Db 2301 LLTRYLRHQPQSW 2313

RESULT 9
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US2003016536A1

QY 301 LLTRYLRHQPQSW 313
Db 1440 LLTRYLRHQPQSW 1452

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/187,319
  FILING DATE: 27-Aug-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 09/523,656
  FILING DATE: 2000-03-10
  APPLICATION NUMBER: US 09/037,601
  FILING DATE: 1998-03-10
  APPLICATION NUMBER: WO PCT/US97/11155
  FILING DATE: 1997-06-26
  APPLICATION NUMBER: US 08/670,707
  FILING DATE: 1996-06-26
ATTORNEY/AGENT INFORMATION:
  NAME: Greenlee, Lorraine L.
  REGISTRATION NUMBER: 27,894
  REFERENCE/DOCKET NUMBER: 75-95K
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 303/499-8080
    TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2332 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: unknown
  MOLECULE TYPE: protein
  HYPOTHEICAL: YES
  ANTI-SENSE: NO
  FRAGMENT TYPE: N-terminal
  ORIGINAL SOURCE:
    ORGANISM: Homo sapiens
    TISSUE TYPE: Liver
  SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

Query Match 100.0%; Score 1666; DB 9; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYSNKCQTPLGMSGHIRDFQITASGOYGQWAPKLARLHYSGSIN 60
Db 2001 LIGELHAGMSTFLVYSNKCQTPLGMSGHIRDFQITASGOYGQWAPKLARLHYSGSIN 2060

QY 61 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 120
Db 2061 AMSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQYRGNST 2120

QY 121 GTLMVFFGNVDSSGKKNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFFGNVDSSGKKNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCMPLGM 2180

QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNARPOVNNPKEWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNARPOVNNPKEWLQVDFQKTMKV 2240

QY 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGQNDSPFPVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGQNDSPFPVNSLDPP 2300

QY 301 LLTRYLRHQPQSW 313
Db 2301 LLTRYLRHQPQSW 2313

RESULT 8
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
GENERAL INFORMATION:
  APPLICANT: Loliar, John S.
  TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
  NUMBER OF SEQUENCES: 40
  CORRESPONDENCE ADDRESS:
    ADDRESS: Greenlee, Winner and Sullivan, P.C.
    STREET: 5370 Manhattan Circle Suite 201
    CITY: Boulder
    STATE: Colorado
    COUNTRY: USA
    ZIP: 80303
COMPUTER READABLE FORM:
```

```

; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-131-510A-2

Query Match      100.0%; Score 1666; DB 14; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 2060

Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

Qy 121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 2180

Qy 181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 2240

Qy 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 2300

Qy 301 LLTRYLRHPQSW 313
Db 2301 LLTRYLRHPQSW 2313

RESULT 10
US-10-445-235-2
; Sequence 2, Application US/10445235
; Publication No. US20040005670A1
; GENERAL INFORMATION:
; APPLICANT: Katherine A. High
; APPLICANT: Rodney M. Canire
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: CHOP0176
; CURRENT APPLICATION NUMBER: US/10/445,235
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: 60/382,486
; PRIOR FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0

Query Match      100.0%; Score 1666; DB 14; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 2060

Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

Qy 121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 2180

Qy 181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 2240

Qy 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 2300

Qy 301 LLTRYLRHPQSW 313
Db 2301 LLTRYLRHPQSW 2313

```

```

; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-445-235-2

Query Match      100.0%; Score 1666; DB 15; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 2060

Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

Qy 121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 2180

Qy 181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNARWPQVNNPKWLQVDFQKTMKV 2240

Qy 241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPP 2300

Qy 301 LLTRYLRHPQSW 313
Db 2301 LLTRYLRHPQSW 2313

RESULT 11
US-10-360-101-229
; Sequence 229, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of factor VIII
; US-10-360-101-229

Query Match      100.0%; Score 1666; DB 15; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 60
Db 2001 LIGHLHAGMSTLFLVYSNKCQTPLGASGHIRDFQITASQYGQWAPKLARLHYSGSIN 2060

Qy 61 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWIKVDLLAPMIHGIKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120

Qy 121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 180
Db 2121 GTLMVFGNVDSSGKINIFNPPIIARIYIRLHPHYSTIRSLRMELMGCDLNSCMPLGM 2180

```

```

QY 181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 2181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 2240
QY 241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 300
DB 2241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 2300
QY 301 LLTRYLRHPQSW 313
DB 2301 LLTRYLRHPQSW 2313

RESULT 12
US-10-239-498A-2
; Sequence 2, Application US/10239498A
; Publication No. US2004002333A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Charlotte
; APPLICANT: Horster, Andrea
; APPLICANT: Schroder, Carola
; APPLICANT: Lehner, Michael
; TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
; TITLE OF INVENTION: Human Cell Lines
; FILE REFERENCE: 80977.0001
; CURRENT APPLICATION NUMBER: US/10/239,498A
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/BP01/03220
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-498A-2

Query Match 100.0%; Score 1666; DB 15; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 60
DB 2001 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 2060
QY 61 AWSTKEPFSWIKVDLLAPMIITHGKIQGARQKFSLLYSQFIIMYSLDGKKWQTYRGNST 120
DB 2061 AWSTKEPFSWIKVDLLAPMIITHGKIQGARQKFSLLYSQFIIMYSLDGKKWQTYRGNST 2120
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
DB 2121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
QY 181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 2181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 2240
QY 241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 300
DB 2241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 2300
QY 301 LLTRYLRHPQSW 313
DB 2301 LLTRYLRHPQSW 2313

RESULT 13
US-10-466-998A-1
; Sequence 1, Application US/10466998A
; Publication No. US20040126856A1
; GENERAL INFORMATION:
; APPLICANT: BAJAJ, S. Paul

```

```

; APPLICANT: FAY, Philip J.
; TITLE OF INVENTION: Factor IXa: Factor VIIa Interaction and Methods
; FILE REFERENCE: 66153-41436
; CURRENT APPLICATION NUMBER: US/10/466,998A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/01724
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 60/263,431
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-998A-1

Query Match 100.0%; Score 1666; DB 16; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 60
DB 2001 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 2060
QY 61 AWSTKEPFSWIKVDLLAPMIITHGKIQGARQKFSLLYSQFIIMYSLDGKKWQTYRGNST 120
DB 2061 AWSTKEPFSWIKVDLLAPMIITHGKIQGARQKFSLLYSQFIIMYSLDGKKWQTYRGNST 2120
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
DB 2121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
QY 181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
DB 2181 ESKAISDAQITASSYFTNMEATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 2240
QY 241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 300
DB 2241 TGVTTQGVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKVFGQNDSTFTPVVNSLDPP 2300
QY 301 LLTRYLRHPQSW 313
DB 2301 LLTRYLRHPQSW 2313

RESULT 14
US-10-721-997A-34
; Sequence 34, Application US/10721997A
; Publication No. US20040197875A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Behring GmbH
; TITLE OF INVENTION: Modified cDNA Factor VIII and its Derivatives
; FILE REFERENCE: 2002/M018-A66
; CURRENT APPLICATION NUMBER: US/10/721,997A
; CURRENT FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-721-997A-34

Query Match 100.0%; Score 1666; DB 16; Length 2332;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 60
DB 2001 LIGEHLHAGMSTFLVYSNKQCTPLGWSGHIRDFOITASQGYGOWAPKLARLHYSGIN 2060

```

Job time : 69.4035 secs

```
Qy 61 AWSTKEPFSWIKVDLLAPMIITHGKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2061 AWSTKEPFSWIKVDLLAPMIITHGKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2120
Qy 121 GTLMVFFGNVDSSGIGKHNIFNPPIIARIYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 180
Db 2121 GTLMVFFGNVDSSGIGKHNIFNPPIIARIYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 2180
Qy 181 ESKAISDAQITASSYFTNMFPATWSPSKARLHLQGRSNARWQPNNPKEWLQVDFQKTMKV 240
Db 2181 ESKAISDAQITASSYFTNMFPATWSPSKARLHLQGRSNARWQPNNPKEWLQVDFQKTMKV 2240
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 2300
Qy 301 LLTRYLRIHPQSW 313
Db 2301 LLTRYLRIHPQSW 2313

RESULT 15
US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth R
; TITLE OF INVENTION: Hoshijima, Masahiko
; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII
; TITLE OF INVENTION: with vesicle vector
; FILE REFERENCE: 6627-Pat170
; CURRENT APPLICATION NUMBER: US/10/132.829
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,314
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-829-4

Query Match 100.0%; Score 1666; DB 14; Length 2351;
Best Local Similarity 100.0%; Pred. No. 1.4e-161;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIGHLHAGNSTLFLVSNKQOTPLGNASGHIRDFQITASQYQWAPKARLHYSGIN 60
Db 2020 LIGHLHAGNSTLFLVSNKQOTPLGNASGHIRDFQITASQYQWAPKARLHYSGIN 2079
Qy 61 AWSTKEPFSWIKVDLLAPMIITHGKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 120
Db 2080 AWSTKEPFSWIKVDLLAPMIITHGKTQARQKFSLSYISQFIIMYSLDGKKWQTYRGNST 2139
Qy 121 GTLMVFFGNVDSSGIGKHNIFNPPIIARIYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 180
Db 2140 GTLMVFFGNVDSSGIGKHNIFNPPIIARIYIRLHPHYHSIRSLRMELMGCDLNSCSMPLGM 2199
Qy 181 ESKAISDAQITASSYFTNMFPATWSPSKARLHLQGRSNARWQPNNPKEWLQVDFQKTMKV 240
Db 2200 ESKAISDAQITASSYFTNMFPATWSPSKARLHLQGRSNARWQPNNPKEWLQVDFQKTMKV 2259
Qy 241 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 300
Db 2260 TGVTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTPVVNSLDPP 2319
Qy 301 LLTRYLRIHPQSW 313
Db 2320 LLTRYLRIHPQSW 2332
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:13:08 ; Search time 21.6931 Seconds
(without alignments)
1388.270 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332

Perfect score: 1666

Sequence: 1 LIGHLHAGMSTFLVSNK.....VNSLDPPLLTRYLRHPSW 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	2351	1 EZHU	coagulation factor
2	1495	89.7	2319	2 A47004	coagulation factor
3	1467	88.1	2133	2 T42763	coagulation factor
4	1006	60.4	216	2 A44258	factor VIII-associ
5	745.5	44.7	2224	1 KFHU5	coagulation factor
6	741.5	44.5	2183	2 T42764	coagulation factor
7	738.5	44.3	2211	1 KFB05	coagulation factor
8	621.5	37.3	427	2 JC4915	agb protein precu
9	618	37.1	463	1 A36479	milk fat globule m
10	606.5	36.4	409	2 T11743	pp47 protein - pig
11	586.5	35.2	401	2 S51138	glycoprotein antig
12	586.5	35.2	427	2 T47211	PAS-6/7 protein pr
13	406	24.4	927	1 JQ0948	A5 antigen precurs
14	381.5	22.9	218	2 A47285	milk fat globule p
15	277.5	16.7	3133	2 S52093	hemocytin - silkw
16	222	13.3	845	2 UC5286	adipocyte transcri
17	211	12.7	1283	2 T13799	neurexin IV - frui
18	202	12.1	737	2 T31349	hypothetical prote
19	191	11.5	737	2 T15615	hypothetical prote
20	185.5	11.1	855	2 S42621	protein-tyrosine k
21	182	10.9	719	2 S51739	transcription repr
22	173.5	10.4	819	2 I48859	tyro 10 receptor k
23	170.5	10.2	913	2 A48280	receptor tyrosine k
24	169.5	10.2	876	2 A49508	protein-tyrosine k
25	167	10.0	910	2 A53137	tyrosine kinase re
26	158	9.5	1385	2 T14158	neurexin IV - mous
27	156	9.4	1381	2 T31083	paranodin - rat
28	119	7.1	791	2 T16031	hypothetical prote
29	113	6.8	149	1 DLD01D	discoidin I chain

30	105.5	6.3	253	1 DLD01A	discoidin I chain
31	105	6.3	253	1 DLD01C	discoidin I chain
32	98.5	5.9	149	2 B03382	discoidin I chain
33	98	5.9	874	4 GNHUER	retrovirus-related
34	94.5	5.7	159	2 AB0486	probable exported
35	91.5	5.5	452	2 A35721	nicotinic acetylch
36	91.5	5.5	1740	2 T43773	hypothetical prote
37	90.5	5.4	1827	2 T34288	hypothetical prote
38	90	5.4	531	1 SYNCYT	tyrosine-tyrosine k
39	89	5.3	815	2 F90951	biotin sulfoxide r
40	89	5.3	815	2 B58000	biotin sulfoxide r
41	88.5	5.3	883	1 JBVLC2	DNA-directed DNA p
42	88	5.3	1070	2 S46755	hypothetical prote
43	87	5.2	4377	2 A55575	ankyrin 3, long sp
44	86	5.2	399	1 JU0348	47.6K protein - Ch
45	86	5.2	455	2 B82658	beta-lactamase XF1

ALIGNMENTS

RESULT 1

EZHU

coagulation factor VIII precursor [validated] - human

N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant compo

C:Species: Homo sapiens (man)

C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004

C:Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445; B42;

R:Gitschier, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

A:Title: Sequence of the exon-containing regions of the human factor VIII gene.

A:Reference number: I54318; MUID:93265012; PMID:1303178

A:Accession: I54318

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1921, 'S', 1923-2351 <RES>

A:Cross-references: UNIPROT:P00451; GB:M88648; NID:g182381; PIDN:AA52420.1; PID:g182383

R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg,

Nature 312, 330-337, 1984

A:Title: Expression of active human factor VIII from recombinant DNA clones.

A:Reference number: A00525; MUID:85061548; PMID:6438526

A:Accession: A00525

A:Molecule type: mRNA

A:Residues: 1-2351 <MOO>

A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179

R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D.;

S. D.N.; Hewick, R.M.

Nature 312, 342-347, 1984

A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.

A:Reference number: I58059; MUID:85061550; PMID:6438528

A:Accession: I58059

A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>

A:Cross-references: GB:X01740; NID:g182802; PIDN:AA52484.1; PID:g182803

R:Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo,

.B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; Noi

DNA 4, 333-349, 1985

A:Title: Characterization of the polypeptide composition of human factor VIII:C and the r

A:Reference number: A23584; MUID:86081164; PMID:3935400

A:Accession: A23584

A:Molecule type: mRNA

A:Residues: 1-2351 <TRU>

A:Cross-references: GB:M14113; NID:g182817; PIDN:AA52485.1; PID:g182818

R:Eaton, D.; Rodriguez, H.; Vehar, G.A.

Biochemistry 25, 505-512, 1986

A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages i

ity.

A:Reference number: A26174; MUID:86159740; PMID:3082357

A:Accession: A26174

A:Molecule type: protein

A:Residues: 20-36,3392-399, 'X', 401-402,1668-1678,1709-1722, 'D', 1723-1725,1741-1755 <EAT>

R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
A:Title: Identification and functional importance of tyrosine sulfate residues within re
A:Reference number: A42348; MUID:92207952; PMID:1554716
A:Accession: A42348
A:Molecule type: protein
A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
A:Experimental source: recombinant material from Chinese hamster ovary cells
R:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
R:Fay, P. J.; Smudzin, T. M.
J. Biol. Chem. 264, 14005-14010, 1989
A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
A:Reference number: A43986; MUID:89340500; PMID:2503509
A:Accession: A43986
A:Molecule type: protein
A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>
R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Rutner, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
A:Reference number: A56109; MUID:91093266; PMID:1898735
A:Contents: annotation; sulfation
R:Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A:Title: Characterization of the human factor VIII gene.
A:Reference number: A56196; MUID:85061547; PMID:6438525
A:Contents: annotation; introns
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
A:Reference number: A56216; MUID:95338127; PMID:7613471
A:Contents: annotation; disulfide bonds
R:Note: 329-Cys, and 2019-Cys were shown to have free sulfhydryls
R:Kjalker, M.; Hedding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
A:Reference number: S63527; MUID:96163459; PMID:8575434
A:Accession: S63527
A:Molecule type: protein
A:Residues: 733-752;753-759 <KJA>
R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baekman, M.; Almstedt, A.; Gray, E.; Sandberg
Eur. J. Biochem. 232, 19-27, 1995
A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
A:Reference number: S66445; MUID:96048024; PMID:7556150
A:Accession: S66445
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1668-1685 <LTN>
C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
A:Gene: GDB:P8C
A:Cross-references: GDB:119124; OMIM:306700
A:Map position: Xq28-Xq28
A:Map position: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
A:Pathway: blood coagulation
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>
F:20-356/Domain: A1 <DAL>
F:23-348/Domain: ferroxidase repeat homology <FO1>
F:392-759/Domain: A2 <DA2>
F:402-730/Domain: ferroxidase repeat homology <FO2>
F:760-1667/Domain: B <DBO>
F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>
F:1709-2038/Domain: A3 <DA3>
F:1716-2038/Domain: ferroxidase repeat homology <FO3>
F:2039-2191/Domain: C1 <DC1>
F:2192-2351/Domain: discoidin I amino-terminal homology <DN1>
F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>
F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #status
F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
F:365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimental
F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:414,426/Binding site: sulfate (Tyr) (covalent) #status experimental
F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1667-1668/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F:2193-2345/Disulfide bonds: #status predicted
Query Match 100.0%; Score 1666; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 2,1e-133;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIGELHAGMSTFLVYVSNKQCPPLGMSGHIRDFQITASGGYQGWAPKPLARLHYSGIN 60
Db 2020 LIGELHAGMSTFLVYVSNKQCPPLGMSGHIRDFQITASGGYQGWAPKPLARLHYSGIN 2079
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTGQARQKFSYISQFIIMYSLDGKWKQYRGNST 120
Db 2080 AWSTKEPFSWIKVDLLAPMIHGIKTGQARQKFSYISQFIIMYSLDGKWKQYRGNST 2139
QY 121 GTLMVFFGNVDSGGIKHNINPPIIARYTLHPTHYSIRTLRMELMGCDLNSCMPLGM 180
Db 2140 GTLMVFFGNVDSGGIKHNINPPIIARYTLHPTHYSIRTLRMELMGCDLNSCMPLGM 2199
QY 181 ESKATSDAQITASSYFTNMATWSPSKARLHLOGRNSAWRPQVNNPKKWLQVDQKTMKV 240
Db 2200 ESKATSDAQITASSYFTNMATWSPSKARLHLOGRNSAWRPQVNNPKKWLQVDQKTMKV 2259
QY 241 TGVTTQGVKSLTSMYKKEFLISSDQGHQWTLFFQNGKVKVFGQNGDSFPPVNSLDPP 300
Db 2260 TGVTTQGVKSLTSMYKKEFLISSDQGHQWTLFFQNGKVKVFGQNGDSFPPVNSLDPP 2319
QY 301 LLTRYLRHPQSW 313
Db 2320 LLTRYLRHPQSW 2332
RESULT 2
A47004
coagulation factor VIII precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1994 #sequence revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A47004
R:Elder, B.; Lakich, D.; Gitschier, J.
Genomics 16, 374-379, 1993
A:Title: Sequence of the murine factor VIII cDNA
A:Reference number: A47004; MUID:93300511; PMID:8314577
A:Accession: A47004
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2319 <ELD>
A:Cross-references: UNIPROT:Q06194; GB:L05573; NID:g192456; PIDN:AAA3785.1; PID:g192457
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:23-349/Domain: ferroxidase repeat homology <FO1>
F:402-730/Domain: ferroxidase repeat homology <FO2>
F:1686-2006/Domain: ferroxidase repeat homology <FO3>
F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>
Query Match 89.7%; Score 1495; DB 2; Length 2319;
Best Local Similarity 89.1%; Pred. No. 7,9e-119;
Matches 279; Conservative 21; Mismatches 21; Indels 0; Gaps 0;
QY 1 LIGELHAGMSTFLVYVSNKQCPPLGMSGHIRDFQITASGGYQGWAPKPLARLHYSGIN 60
Db 1988 LIGELHAGMSTFLVYVSNKQCPPLGMSGHIRDFQITASGGYQGWAPKPLARLHYSGIN 2047
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTGQARQKFSYISQFIIMYSLDGKWKQYRGNST 120


```
Db 2048 AMSTKEPFSWIKVDLLAPMIVHGKIQGARKFSSLYISQFIIMVSLDGGKWLVSQGNST 2107
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLM 180
Db 2108 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLM 2167
QY 181 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
Db 2168 ESKVLSDSQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 2227
QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPP 300
Db 2228 TGIITQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPP 2287
QY 301 LLTRYLRHQPQSW 313
Db 2288 LLTRYLRHQPQSW 2300

RESULT 3
T42763
coagulation factor VIII precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42763
R:Collar, P.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z22269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2133 <LOL>
A:Cross-references: UNIPROT:P12263; EMBL:U49517; NID:G1511633; PID:G1511634; PIDN:AA8067
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 88.1%; Score 1467; DB 2; Length 2133;
Best Local Similarity 86.3%; Pred. No. 1.7e-116;
Matches 270; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 1 LIGELHAGMSTFLVYGNKQOTPLGMASGHRDFOITASGOYGOWAPKLARLHYSGSIN 60
Db 1802 LIGELHAGMSTFLVYGNKQOTPLGMASGHRDFOITASGOYGOWAPKLARLHYSGSIN 1861
QY 61 AMSTKEPFSWIKVDLLAPMIVHGKIQGARKFSSLYISQFIIMVSLDGGKWLVSQGNST 120
Db 1862 AMSTKDPHSWIKVDLLAPMIVHGKIQGARKFSSLYISQFIIMVSLDGRNWSYRGNST 1921
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLM 180
Db 1922 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLM 1981
QY 181 ESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 240
Db 1982 QNKALSDSQITASSLSNIFATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKV 2041
QY 241 TGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPP 300
Db 2042 TGIITQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPP 2101
QY 301 LLTRYLRHQPQSW 313
Db 2102 LLTRYLRHQPQSW 2114

factor VIII-associated gene B hypothetical protein - human

RESULT 4
A44258
coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
A:Reference number: Z22269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2133 <LOL>
A:Cross-references: UNIPROT:P12263; EMBL:U49517; NID:G1511633; PID:G1511634; PIDN:AA8067
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>
```

```
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: A44258
R:Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-599, 1992
A:Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; MUID:93052386; PMID:1427887
A:Accession: A44258
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-216 <LEV>
A:Cross-references: UNIPROT:Q14286; GB:M90707; NID:G182316; PIDN:AA58466.1; PID:G182317
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F:1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
F:57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 60.4%; Score 1006; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.8e-78;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 VFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLMESKA 184
Db 9 VFFGNVDSSGKHNIFNPPIIARYIRLHPTHTYSIRSLRMELMGCDLNSCSMPGLMESKA 68
QY 185 ISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKVGT 244
Db 69 ISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPOVNNPKWLQVDFQKTMKVGT 128
QY 245 TQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPPLTR 304
Db 129 TQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNGDSFTPVVNSLDPPLTR 188

RESULT 5
KPFUS
coagulation factor V precursor [validated] - human
N:Alternate names: coagulation labile factor; proaccelerin
C:Species: Homo sapiens (man)
C:Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 08-Dec-2000
C:Accession: A56172; A42344; A28028; A27498; A25897
R:Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A:Title: Structure of the gene for human coagulation factor V.
A:Reference number: A42344; MUID:92232668; PMID:1567832
A:Accession: A56172
A:Molecule type: DNA
A:Residues: 1-2224 <CRI>
A:Cross-references: GB:J05368
A:Accession: A42344
A:Molecule type: DNA
A:Residues: 48-58179-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-542;
2070;2111-2172-2181 <CR2>
R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, J.J.; Aldape, R.A.; Hewick, R.M.; Kaufm
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A:Title: Complete cDNA and derived amino acid sequence of human factor V.
A:Reference number: A28028; MUID:87260886; PMID:3110773
A:Accession: A28028
A:Molecule type: mRNA
A:Residues: 1-857; 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <JEN>
A:Cross-references: GB:M16967
A>Note: parts of this sequence, including the amino end of the mature protein, were deter
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum
A:Reference number: A27498; MUID:88107560; PMID:2827731
A:Accession: A27498
A:Molecule type: mRNA
A:Residues: 1-1284, 'I', 1286-1600 <KAN>
A:Cross-references: GB:M17785
```

A;Note: parts of this sequence were determined by protein sequencing
R;Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A;Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog
A;Reference number: A25897; MUID:86313665; PMID:3092220
A;Accession: A25897
A;Molecule type: mRNA
A;Residues: 1188-1215, 1315-2224 <KA2>
A;Cross-references: GB:M14335
A;Note: parts of this sequence were determined by protein sequencing
R;Keller, F.G.; Ortel, T.L.; Quinn-Allen, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A;Title: Thrombin-catalyzed activation of recombinant human factor V.
A;Reference number: A56139; MUID:95210278; PMID:7696276
A;Contents: annotation; thrombin cleavage sites
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Genetics:
A;Gene: GDB:F5
A;Cross-references: GDB:119896; OMIM:227400
A;Map position: 1q23-1q23
A;Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1; 65
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prot
C;Superfamily: coagulation factor V; discoidein I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-737/Product: coagulation factor V #status predicted <MAT>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-691/Domain: A2 <DA2>
F;351-684/Domain: ferroxidase repeat homology <FO2>
F;692-1573/Domain: B <DOB>
F;1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
F;1574-1905/Domain: A3 <DA3>
F;1581-1905/Domain: ferroxidase repeat homology <FO3>
F;1667-1765/Region: phospholipid binding #status predicted
F;1906-2061/Domain: Cl <DC1>
F;2065-2221/Domain: discoidein I amino-terminal homology <DN1>
F;2065-2221/Domain: C2 <DC2>
F;53-55-239-297-460-468-554-741-752-760-776-782-821-938-977-1074-1083-1103-1106-1479-149
F;167-193-248-329-500-526-603-684-1725-1751-1907-2061-2066-2221/Disulfide bonds: #statu
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363-693/Binding site: sulfate (tyr) (covalent) #status predicted
F;376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;382-1338/Binding site: carbohydrate (Asn) (covalent) #status absent
F;534-535/Cleavage site: Arg-Gly (protein C) #status predicted
F;737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experimental
F;1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 44.7%; Score 745.5; DB 1; Length 2224;
Best Local Similarity 44.9%; Pred. No. 6.6e-55;
Matches 144; Conservative 65; Mismatches 103; Indels 9; Gaps 4;
QY 2 IGHHLHAGMSTLFLVYSNCKOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSINA 61
DB 1888 VGENQACQWTFLLMDRDCRMPGLSTGLISDSQIKASEFLGYEPLRLNNGGSYNA 1947
QY 62 WSTKE----PFS---WIKVDLLAPMIHGIKTQAGKQKFSLSYISQFIIMYSLDGKKQTY 115
DB 1948 WSEVKLAELAEFASKPWQVDMQKEVITIGTQAGAKHYLSCTTFEYVAYSQINQWIP 2007
QY 116 RGNSTGLMVFNGVDSGKIHNFPIIARVIRLPHYTHYSIRSTLRMLMGCDLNSCS 175
DB 2008 KGNSTRNVYFNGSDASTIKENQDFPIVARIYIRISPTAVNRPTLRLELQGCVEVNGS 2067
QY 176 MFLGMSKAISDAQITASSYFTNNFAT--WSPSKARLHLOGRSNARPOVNNPKWLQVDF 234
DB 2068 TPLGMENGIENKQITASSPFKSWGDIWEPFARLNAQGRVNAQKANNKQWLQVDF 2127

QY 235 QKTMKVTGVTGQVKSLLTSMYVKEFLISSQDCHQWTLFFQNGKV--KVFGNODSFTP 292
DB 2128 LTKKKITAITGCKSLSEMYKSYTIHYSEQGVKPKYRLKSSMVDKIFEGNTIKGH 2187
QY 293 VNNSLDPPLLTRYLRHPQSW 313
DB 2188 VKNFFNPPIIRFIRIPKTW 2208
RESULT 6
T42764
coagulation factor V - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42764
R;Yang, T.L.; Cui, J.; Rehmtulla, A.; Yang, A.; Mousalli, M.; Kaufman, R.J.; Ginsburg,
Blood 91, 4593-4599, 1998
A;Title: The structure and function of murine factor V and its inactivation by protein C.
A;Reference number: 222270; MUID:98282202; PMID:9616155
A;Accession: T42764
A;Status: preliminary; translated from GB/EMBL/DBU
A;Molecule type: mRNA
A;Residues: 1-2183 <YAN>
A;Cross-references: UNIPROT:O88783; EMBL:U52925; NID:G3219690; PID:G3219691; PIDN:AAC995;
C;Function:
A;Pathway: blood coagulation
A;Superfamily: coagulation factor V; discoidein I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma; f
F;350-682/Domain: ferroxidase repeat homology <FOX1>
F;1541-1864/Domain: ferroxidase repeat homology <FOX2>
Query Match 44.5%; Score 741.5; DB 2; Length 2183;
Best Local Similarity 45.2%; Pred. No. 1.4e-54;
Matches 145; Conservative 60; Mismatches 107; Indels 9; Gaps 4;
QY 2 IGHHLHAGMSTLFLVYSNCKOTPLGASGHIRDFQITASQYQGWAPKLARLHYSGSINA 61
DB 1847 VGENQACQWTFLLMDRDCRMPGLSTGLISDSQIKASEYLYTWEPLRLNNGSINA 1906
QY 62 WSTKE----PFS---WIKVDLLAPMIHGIKTQAGKQKFSLSYISQFIIMYSLDGKKQTY 115
DB 1907 WSEKTAALDFPKPIQVDMQKEVVTGIGTQAGAKHYLSCTTFEYVAYSQDQTNWQIF 1966
QY 116 RGNSTGLMVFNGVDSGKIHNFPIIARVIRLPHYTHYSIRSTLRMLMGCDLNSCS 175
DB 1967 RGSKGSVMYFTGNSDGTIKENRDLDPPIVARIYIRIHPKTSYNRPTLRLELQGCVEVNGS 2026
QY 176 MFLGMSKAISDAQITASSYFTNNFAT--WSPSKARLHLOGRSNARPOVNNPKWLQVDF 234
DB 2027 TPLGEDGRIQDKQITASSPFKSWGDIWEPFARLNAQGRVNAQKANNKQWLQVDF 2086
QY 235 QKTMKVTGVTGQVKSLLTSMYVKEFLISSQDCHQWTLFFQNGKV--KVFGNODSFTP 292
DB 2087 LTKKKITAITGCKSLSEMYKSYTIHYSEQGVKPKYRLKSSMVDKIFEGNTIKGH 2146
QY 293 VNNSLDPPLLTRYLRHPQSW 313
DB 2147 MKNFFNPPIIRFIRIPKTW 2167
RESULT 7
KFBO5
coagulation factor V precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C;Accession: A42580; A36497
R;Guinto, E.R.; Bgmon, C.T.; Mann, K.G.; MacGillivray, R.T.
J. Biol. Chem. 267, 2971-2978, 1992
A;Title: The complete cDNA sequence of bovine coagulation factor V.
A;Reference number: A42580; MUID:92147638; PMID:1737753
A;Accession: A42580
A;Molecule type: mRNA

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-463 <STU>
A:Cross-references: UNIPROT:P21956; GB:M38337; NID:ig199142; PIDN:AAA19534.1; PID:ig199143
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
C:Keywords: membrane protein
F:128-60/Domain: EGF homology <EG1>
F:147-303/Domain: discoidin I amino-terminal homology <DN1>
F:307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 37.1%; Score 618; DB 1; Length 463;
Best Local Similarity 41.7%; Pred. No. 6.4e-45;
Matches 128; Conservative 59; Mismatches 108; Indels 12; Gaps 5;

QY 18 SNKQTPGLMGASHIRDPQITASQY-----GQWAPKLARLHYSGSINAW--STKEPFS 69
DB 145 ASRCSQTQLMGEGAIADQSISASVYMGFQWGLQWGPGLARLYRTGIVNAHNASYDSKP 204

QY 70 WIKVDLLAPMIHIGIKTQAGRKSSLYISQFIIMYSLDGKKWQYRGNSTGTLMVFFGN 129
DB 205 WIQVNLRRKRVSGWMTQASRAGAEYLTFTKVAISLDGRKFEFIQDESQD-KEFLGN 263

QY 130 VDSGGIKHNFPIIARIYIRLPHYSIRSTRLMELMGCDLNSCMPLGMSKASDAQ 189
DB 264 LDNNSLKVNMFTLEAQYIRLYPVSVCHRGCTLRFEGLGCELHGLLEPLGLKNNFTPDQ 323

QY 190 ITASSYFT--NMFA-TWSPSKARLHLQGRSNARPPQVNNPKWLQVDFQTKMKVTGVTQ 246
DB 324 MSASSSYKTNLRAFQWYFHLGRLDNQGKINAWTAQNSAKEWLQVLDLGTQRTQGTIITQ 383

QY 247 GVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVQFQGNQDSFTPVVNSLDPLLRYL 306
DB 384 GARDPCHIQVVSYSKVAHSDGQWTVYVEQSGSKVQFQGNLDNNSHKKNIKEKPFMAYV 443

QY 307 RIHPQSW 313
DB 444 RVLPSW 450

RESULT 10
P47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11743
R:Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtko, J.; Matsuda, T.; Toepfe
Biol. Reprod. 58, 1057-1064, 1998
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated z
A:Reference number: Z17325; MUID:98206817; PMID:9546740
A:Accession: T11743
A>Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-409 <ENS>
A:Cross-references: UNIPROT:P79385; EMBL:Y11683; NID:g2652927; PIDN:CAA72379.1; PID:g265
A:Experimental source: testis
C:Function:
A:Description: may be involved in membrane remodeling and/or function as a zona pellucid
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F:6-40/Domain: EGF homology <EG>

Query Match 36.4%; Score 606.5; DB 2; Length 409;
Best Local Similarity 40.8%; Pred. No. 5.2e-44;
Matches 125; Conservative 55; Mismatches 113; Indels 13; Gaps 4;

QY 21 CQTPGLMGASHIRDPQITASQY-----GQWAPKLARLHYSGSINAW--STKEPFSWIK 72
DB 91 CNAPLGMETGAIDFQISASSMHLGFMGLQWAPELARLHRAIGIVNAHNASYDRNPW 150

QY 73 VDLAPMIHIGIKTQAGRKSSLYISQFIIMYSLDGKKWQYRGNSTGTLMVFFGNVDS 132
DB 151 VNLRRMRTVGTWQASRAGSAEYMKTFKVAYSTDGRKFQIQGAERSGDKIFMGNLN 210

QY 133 SGIKHNFPIIARIYIRLPHYSIRSTRLMELMGCDLNSCMPLGMSKASDAQITA 192
DB 211 SGUKVNLFEVLEQVRLVPLIICHRGCTLRFEGLGCELGLSCAEPGLKNDTIPNKQITA 270

QY 193 SSYFTN---MPATSPSKARLHLQGRSNARPPQVNNPKWLQVDFQTKMKVTGVTQGVK 249
DB 271 SSFRTWGLSAFSAFYPFYARLDNQGKFNAMTAQNSASEWLQIDLGSQRRVTGIIQTGAR 330

QY 250 SLTSMYKVEFLISSQDGHQWTLFFQNGKV--KVFQGNQDSFTPVVNSLDPLLRYL 307
DB 331 DFGHIQVAAKYVAYSDGVSQWTEYRDQGALEGIKIPGNLDNNSHKKNMFTETPLTRFVR 390

QY 308 IHQPSW 313
DB 391 ILFVAV 396

RESULT 11
S65138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
A:Alternate names: glycoprotein component 16/major fat-globule membrane protein/WFG-E8 hc
N:Species: Bos primigenius taurus (cattle)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C:Accession: S65138; G48394
R:Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A:Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal ant
A:Reference number: S65138; MUID:96125736; PMID:8541316
A:Accession: S65138
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-401 <AOX>
R:Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A:Reference number: A48394; MUID:93250576; PMID:8485470
A:Accession: G48394
A>Status: preliminary
A:Molecule type: protein
A:Residues: 207-220 <MAT>
A:Experimental source: milk
A:Note: sequence extracted from NCBI backbone (NCBI:P131457)
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolog
C:Keywords: glycoprotein
F:1-32/Domain: EGF homology (fragment) <EG1>
F:40-79/Domain: EGF homology <EG2>
F:82-239/Domain: discoidin I amino-terminal homology <DN1>
F:243-401/Domain: discoidin I amino-terminal homology <DN2>

Query Match 35.2%; Score 586.5; DB 2; Length 401;
Best Local Similarity 39.3%; Pred. No. 2.6e-42;
Matches 121; Conservative 58; Mismatches 112; Indels 17; Gaps 5;

QY 21 CQTPGLMGASHIRDPQITASQY-----GQWAPKLARLHYSGSINAW--STKEPFSW 70
DB 83 CTSPLGMQGTGAIDFQISASSMHLGFMGLQWAPELARLHRTGIVNAHNASYDRNPW 140

QY 71 IKVDLLAPMIHIGIKTQAGRKSSLYISQFIIMYSLDGKKWQYRGNSTGTLMVFFGNV 130
DB 141 IQVNLRRMRTVGTWQASRAGSAEYLTFTKVAYSTDGRKFQIQVAGRSGDKIFGNV 200

QY 131 DSGIKHNFPIIARIYIRLPHYSIRSTRLMELMGCDLNSCMPLGMSKASDAQI 190
DB 201 NNSGLKINLFDTPLETOQVRLVPLIICHRGCTLRFEGLGCELGCTEPGLKNDTIPNKQI 260

QY 191 TASSYFTN---MFATSPSKARLHLQGRSNARPPQVNNPKWLQVDFQTKMKVTGVTQ 247
DB 261 TASSYKTKWGLSAFSAFYPFYARLDNQGKFNAMTAQNSASEWLQIDLGSQRRVTGIIQT 320

QY 248 VASLLTSMYKVEFLISSQDGHQWTLFFQNG--KVKVQGNQDSFTPVVNSLDPLLRYL 305
DB 321 ARDFGHIQVAAKYVAYSDGVSQWTEYRDQGALEGIKIPGNLDNNSHKKNMFTETPLTRFVR 380

QY 306 LRIHPQSW 313
 Db 381 VRIQPAW 388

RESULT 12

S74211

PAS-6/7 protein precursor - bovine
 N:Alternate names: glycoprotein component 16/major fat-globule membrane protein(MFG-E8 h
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 04-Dec-1997 #sequence revision 12-Dec-1997 #text_change 09-Jul-2004
 C:Accession: S74211; S78114; S24181; S65138; G48394
 R:Hvarregaard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
 Eur. J. Biochem. 240, 628-636, 1996
 A:Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat glob
 A:Reference number: S74211; MUID:97008954; PMID:8856064
 A:Accession: S74211
 A:Molecule type: mRNA
 A:Residues: 1-427 <HVA>
 A:Cross-references: UNIPROT:Q95114; EMBL:X91895; NID:g1632778; PIDN:CAAG2997.1; PID:g163
 A:Accession: S78114
 A:Molecule type: protein
 A:Residues: 19-85;96-110;140-165;174-216;221-232;248-277;285-293;309-337;339-420;425-427
 R:Kim, D.H.; Kanno, C.; Mizokami, Y.
 Biochim. Biophys. Acta 1122, 203-211, 1992
 A:Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, from
 A:Reference number: S23926; MUID:92353107; PMID:1643094
 A:Accession: S24181
 A:Molecule type: protein
 A:Residues: 383-394 <KIM>
 R:Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
 Biochim. Biophys. Acta 1245, 385-391, 1995
 A:Title: Molecular cloning of glycoprotein antigens MGPS7/53 recognized by monoclonal an
 A:Reference number: S65138; MUID:96125736; PMID:8541316
 A:Accession: S65138
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 27-427 <AOK>
 R:Maicher, I.H.; Banghart, L.R.; Lane, W.S.
 Biochem. Mol. Biol. Int. 29, 545-554, 1993
 A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
 II-like sequences.
 A:Reference number: A48394; MUID:93250576; PMID:8485470
 A:Accession: G48394
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 233-246 <MAT>
 A:Experimental source: milk
 A:Note: sequence extracted from NCBI backbone (NCBIP:131457)
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 C:Keywords: blocked amino end; disulfide bond; glycoprotein; milk
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-427/Product: PAS-6/7 protein #status experimental <MAT>
 F:24-58/Domain: EGF homology <EG1>
 F:66-105/Domain: EGF homology <EG2>
 F:108-265/Domain: discoidin I amino-terminal homology <DN1>
 F:269-427/Domain: discoidin I amino-terminal homology <DN2>
 F:24-35, 29-47, 49-58, 66-77, 71-94, 96-105/Disulfide bonds: #status predicted
 F:27/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:34/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F:59, 227/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:109-265, 252-256, 270-427/Disulfide bonds: #status experimental

Query Match 35.2%; Score 586.5; DB 2; Length 427;
 Best Local Similarity 39.3%; Pred. No. 2.8e-42;
 Matches 121; Conservative 58; Mismatches 112; Indels 17; Gaps 5;

QY 21 COTPLGASGHIRDFQITASGOY-----GQWAPKLARLHYSGSINAWST---KEPFSW 70
 Db 109 CTSPLGQWGTGAIASQISASSWHGLGQWAPKLARLHQTGIWAWTSNGYKNP--W 166

QY 71 IKVDLLAPMIHIGIKTGAROKRQKFSLSYISQFIIMYSLDGKKWQTYRGNSTGLMVFFGNV 130

Db 167 IQVNLRRKMWVTGVVTVGASRAGSAEYLKTFKVAYSTDGRQFQIQVAGRSGDKIFIGNV 226
 QY 131 DSSGIGKHNIFNPPIIARYIRLPHYSIRSTRMELMGCDLNSCMPLGMESKAISDAQI 190
 Db 227 NNSGLKINLFDTPLETOYVRLVPIICHRCGCTLRPELLGCELNGCTEPLGLKNDTIPNKQI 286
 QY 191 TASSYFTN---MFATWSPSKARLHLQGRSNARWQVNNPKEWLQVDFQKTKWKTQVTTQ 247
 Db 287 TASSYKTTWGLSAPSWFPYARLDNQCCKFNATQATNSASEWLQTLDLGSRKRVGTIITQ 346
 QY 248 VKSLTSMYKFEFLISSQDGHQWTLFPQNG--KVKVFQGNQDSFTPVVNSLDLPPLLTRY 305
 Db 347 ARDFGHTQYVAARYAVGDDGVWTEYKDPGASESKIFPGNNDNNHKKNFETFPQARF 406
 QY 306 LRIHPQSW 313
 Db 407 VRIQPAW 414

RESULT 13

QJ0948

A5 antigen precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: JH0466; JQ0948
 R;Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
 Neuron 7, 295-307, 1991
 A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homologie
 A:Reference number: JH0466; MUID:91337458; PMID:1908252
 A:Accession: JH0466
 A:Molecule type: mRNA
 A:Residues: 1-927 <TAK>
 A:Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BAA01260.1; PID:g222963
 A:Experimental source: tadpole, brain
 A:Note: this protein has motifs homologous to complement components C1r and C1s and to c
 C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal rec
 C:Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-termi
 C:Keywords: duplication; glycoprotein; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-927/Product: A5 antigen #status predicted <A5A>
 F:27-138/Domain: C1r/C1s repeat homology <C1R1>
 F:147-262/Domain: C1r/C1s repeat homology <C1R2>
 F:274-424/Domain: discoidin I amino-terminal homology <DN1>
 F:430-584/Domain: discoidin I amino-terminal homology <DN2>
 F:646-812/Domain: MAM homology <MAM>
 F:861-883/Domain: transmembrane #status predicted <TM>
 F:150, 261, 300, 523, 844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.4%; Score 406; DB 1; Length 927;
 Best Local Similarity 34.4%; Pred. No. 1.9e-26;
 Matches 108; Conservative 54; Mismatches 126; Indels 26; Gaps 12;

QY 15 LVYSN-----KCQTPLGASGHIRDFQITASGOY-GQWAPKLARLHYSGSINAWSTKEPP 68
 Db 264 VVQSNTDEDFQCKALGMSGEIHFQDQISVSSQSMNSAERSRLNYVE--NGWTPGDT 321
 QY 69 --SWIKVDLLAPMIHIGIKTQGA--ROKFSLSYISQFIIMYSLDGKKWQTYR-GNSTGT 123
 Db 322 VKEWIQVDLENLRFVSGIGTQGALSKETKKYFKVSKYVDISSNGEDWITLKDGNKH--- 378
 QY 124 MVFFGNVDSGIGKHNIFNPPIIARYIRLPHYSIRSTRMELMGCDLNS--CSMPLGME 181
 Db 379 LVFTGNTDATDVVYRPFSPKPVITRFVRLRPVTVWENGISLRFELYGKITDYPCSRMLGMV 438
 QY 182 SKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWQVNNP--KEWLVDFQKTMK 239
 Db 439 SGLISDQITASS---QVDRNWPVELARLVTSRSGWALPPSNTHPTYKLEWLQIDLAEEKI 495
 QY 240 VTGVTTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNG--GKVKVFQGNQDSFTPVVNSL 297
 Db 496 VRGVIIQGGKHKNKVMRPFKIGYSNNGTEWEMIMDSSKNKPKTFEGNTNIDYTPELKTF 555

QY 298 DPPLLTRYLRHPQ 311
 Db 556 -AHITTFIRIIEP 568

RESULT 14
 A47285
 milk fat globule protein - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 C:Accession: A47285
 R:Larocca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
 Cancer Res. 51, 4994-4998, 1991
 A:Title: A M-r 46,000 human milk fat globule protein that is highly expressed in human b
 A:Reference number: A47285; MUID:91371351; PMID:1909932
 A:Accession: A47285
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-218 <LAR>
 A:Cross-references: UNIPROT:Q08431; GB:S56151; NID:G235396; PIDN:AA819771.1; PID:G235397
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 F:1-56/Domain: discoidin I amino-terminal homology (fragment) <DNI>
 F:60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match 22.9%; Score 381.5; DB 2; Length 218;
 Best Local Similarity 39.2%; Pred. No. 3.6e-25;
 Matches 76; Conservative 40; Mismatches 71; Indels 7; Gaps 3;

QY 126 PFGNVDSSGIGKHNIPPIIARYIRLHPHTYISIRSLRMELMGCDLNSCMPLGWSKAI 185
 Db 13 FVGWKNKNAVHVNLPETVEAQYVRLYPTSGHTACTLRFELGCELNGCANPLGLKNSI 72
 QY 186 SDAQITASSYF---TNMFATWSPSKARLHLOGSNARPOVNNPEWLOVDFOKTMKVT 241
 Db 73 PDKQITASSYKTGHLHF-SNFSYARLDKQGNFRAWAGSYGNDQWLQVDLGSSKEVT 131
 QY 242 GVTTOGVKSLTSMYKFEFLISSODGHWTLFF--ONCKVKVFGNQDSFTPVVNSLD 299
 Db 132 GIITQGARFGVQVFSYKWAYNSDANWTEQDPTGSSKIFFPGNWDNHSKKNLFT 191
 QY 300 PLTRYLRHPQSW 313
 Db 192 PILARYVRLPVAW 205

RESULT 15
 S52093
 hemocytin - silkworm
 N:Alternate names: humoral lectin
 C:Species: Bombyx mori (silkworm)
 C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 12-Jul-2004
 C:Accession: S52093; S70920
 R:Kotani, E.; Yamakawa, M.; Iwamoto, S.; Tashiro, M.; Mori, H.; Sumida, M.; Matsubara, F
 Biochim. Biophys. Acta 1260, 245-258, 1995
 A:Title: Cloning and expression of the gene of hemocytin, an insect humoral lectin which
 A:Reference number: S52093; MUID:95178544; PMID:7873598
 A:Accession: S52093
 A:Molecule type: mRNA
 A:Residues: 1-3133 <KOT>
 A:Cross-references: UNIPROT:P98092; EMBL:D29738
 R:Mori, H.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S70920
 A:Accession: S70920
 A:Molecule type: mRNA
 A:Residues: 1-1566, 'S', 1568-3133 <MOR>
 A:Cross-references: EMBL:D29738; NID:G474967; PID:G664884
 C:Superfamily: hemocytin/hemocytin

Query Match 16.7%; Score 277.5; DB 2; Length 3133;
 Best Local Similarity 25.9%; Pred. No. 8.8e-15;
 Matches 82; Conservative 61; Mismatches 124; Indels 49; Gaps 13;

QY 25 LGMASGHIRDFOITASGOVGO-WAPKRLARLH---YSGSINAWSTK--EPFWSIKVDLLAP 78
 Db 948 LVMGDEPLPDATFASSESEIFAPHNARLNRGPTNSGAGSNPKVNDKQYIQVELPRR 1007
 QY 79 MIHGIKTGQARQKFSLSYISOFIMYSLDGKKWQTYRGNSGTGLMVFGNVDSGSIKHN 138
 Db 1008 EPIYGVVLGS--PIFDQYVTSYSEIMYGDNDONFVSVDG-PDGKPKIFAGPIDNHPVKQ 1064
 QY 139 IFNPPIIARYIRLHPHTYISIRSLRMELMGCDLNSCMPLGWSKAI 182
 Db 1065 MISPPIEAKVVRIRLPTWHDEISLRLEIIGCABPLTFTTSETSESPLOCTEPLGLIG 1124
 QY 183 K-AISDAQITASSYFTNMFATWSPSKARLHLOGSNARPOVNNPEWLOVDFOKTMKVT 241
 Db 1125 ELPLENIQVSSN-----SEKDYLSING-NRGWKPLYNTF-GWVMFDFGTGPRNIT 1172
 QY 242 GVTTOGVK-SLLTS---MYKFEFLISSODGHWTLFFQNGKVKVFGNQDSFTPVVNSL 297
 Db 1173 GILTKGNDGWVTSYKVLVTSDFETEN-----PVIDKDGKEKIFPANFDGIVSVTNEF 1225
 QY 298 DPPLLTRYLRHPQSW 313
 Db 1226 HPIRARYLKVLPQKW 1241

Search completed: April 20, 2005, 17:21:42
 Job time : 23.6931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:11:58 ; Search time 78.25 Seconds
(without alignments)
2048.317 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332

Perfect score: 1666

Sequence: 1 LIGHLHAGMSTFLVYSNK.....VNSLDPLRLYLRHPQSW 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	2351	1	FA8_HUMAN
2	1522	91.4	2343	2	O18806
3	1517	91.1	2343	2	O62730
4	1495	89.7	2319	1	FA8_MOUSE
5	1467	88.1	2133	1	FA8_PIG
6	1264	75.9	2258	2	Q7TN96
7	1146	68.8	1377	2	Q804X3
8	1006	60.4	216	2	Q14286
9	846	50.8	1639	2	Q804W6
10	782	46.9	745	2	Q804X4
11	771.5	46.3	1460	2	Q7SZN0
12	745.5	44.7	2224	1	FA5_HUMAN
13	745.5	44.7	2224	1	Q6UP06
14	743.5	44.6	2258	1	FA5_PIG
15	741.5	44.5	2183	2	O88783
16	738.5	44.3	2211	1	FA5_BOVIN
17	721	43.3	407	2	Q8AYE0
18	715.5	42.9	1802	2	Q804W5
19	685	41.1	2119	2	Q90X47
20	639.5	38.4	481	2	Q6RUW2
21	625.5	37.5	480	1	ED13_HUMAN
22	625.5	37.5	480	2	Q8N6T0
23	625	37.5	463	1	MFGM_MOUSE
24	621.5	37.3	427	1	MFGM_RAT
25	618.5	37.1	470	2	Q8C4U8
26	618.5	37.1	480	1	ED13_MOUSE
27	618.5	37.1	480	2	Q8CBF7
28	606.5	36.4	409	1	MFGM_PIG
29	586.5	35.2	427	1	MFGM_BOVIN
30	584.5	35.1	475	2	Q6P1V9
31	571.5	34.3	475	2	Q68EW5

```

32 544 32.7 387 1 MFGM_HUMAN Q08431 homo sapien
33 520.5 31.2 2102 2 Q7TPK2 Q7tpk2 rattus norv
34 512 30.7 312 2 Q7Z3D2 Q7z3d2 homo sapien
35 506 30.4 363 2 Q77718 Q77718 equus caball
36 466.5 28.0 384 2 Q8C8K0 Q8c8k0 mus musculus
37 444 26.7 335 2 Q9BTL9 Q9btl9 homo sapien
38 421 25.3 901 2 Q9H2D5 Q9h2d5 homo sapien
39 421 25.3 901 2 Q9H2E4 Q9h2e4 homo sapien
40 421 25.3 906 2 Q9H2D4 Q9h2d4 homo sapien
41 421 25.3 906 2 Q9H2E3 Q9h2e3 homo sapien
42 421 25.3 925 1 NRP2_RAT Q35276 rattus norv
43 421 25.3 931 1 NRP2_HUMAN Q60462 homo sapien
44 421 25.3 931 2 Q7Z3T9 Q7z3t9 homo sapien
45 419.5 25.2 926 2 Q8QZY7 Q8qzy7 mus musculus

```

ALIGNMENTS

```

RESULT 1
FA8_HUMAN
ID FA8_HUMAN STANDARD; PRT; 2351 AA.
AC P00451;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component)
DE (Antihemophilic factor) (AHF).
GN Name=F8; Synonyms=F8C;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Masiarz F.R., Merryweather J.P., Najarian R.,
RA Pahl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA Nordfang O., Ezban M.;
RA "Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA.";
RL DNA 4:333-349 (1985).
RP [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,
RA Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RA "Expression of active human factor VIII from recombinant DNA clones.";
RL Nature 312:330-337 (1984).
RP [3]
RX SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA Hewick R.M.;
RA "Molecular cloning of a cDNA encoding human antihemophilic factor.";
RL Nature 312:342-347 (1984).
RP [4]
RX SEQUENCE FROM N.A.
RX MEDLINE=93265012; PubMed=1303178;
RA Gitschier J., Wood W.I.;
RA "Sequence of the exon-containing regions of the human factor VIII
RL gene.";
RL Hum. Mol. Genet. 1:199-200 (1992).
RP [5]
RX SEQUENCE OF 2064-2070 FROM N.A.
RA de Water N.S., Williams R., Browett P.J.;
RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

```

[6] RX SULFATION OF TVR-1699.
 RA MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyrl680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor."; Chem. 266:740-746 (1991).
 RL J. Biol. Chem. 266:740-746 (1991).
 RP [7]
 RX SULFATION.
 RA MEDLINE=92207952; PubMed=1554716;
 RA Pittman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate residues
 RT within recombinant factor VIII."; Biochemistry 31:3315-3325 (1992).
 RL Biochemistry 31:3315-3325 (1992).
 RP [8]
 RX SULFATION SITES TVR-737; TVR-738 AND TVR 742, AND DISULFIDE BONDS.
 RA PubMed=7613471;
 RA McMullen B.A., Fujikawa K., Davie E.W., Hedner U., Ezban M.;
 RT "Locations of disulfide bonds and free cysteines in the heavy and
 RT light chains of recombinant human factor VIII (antihemophilic factor
 RT A)"; Protein Sci. 4:740-746 (1995).
 RL Protein Sci. 4:740-746 (1995).
 RP [9]
 RX STRUCTURE BY NMR OF 2322-2343.
 RA MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Saleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy."; Biochemistry 34:3022-3031 (1995).
 RL Biochemistry 34:3022-3031 (1995).
 RP [10]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RA MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A."; Ann. N. Y. Acad. Sci. 614:89-96 (1991).
 RL Ann. N. Y. Acad. Sci. 614:89-96 (1991).
 RP [11]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RA MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A."; Blood 73:1-12 (1989).
 RL Blood 73:1-12 (1989).
 RP [12]
 RX REVIEW ON MOLECULAR BASIS OF HEMA.
 RA MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A."; Hum. Mutat. 5:1-22 (1995).
 RL Hum. Mutat. 5:1-22 (1995).
 RP [13]
 RX VARIANT HEMA GLN-2326.
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophilic."; Science 232:1415-1416 (1986).
 RL Science 232:1415-1416 (1986).
 RP [14]
 RX VARIANT HEMA PRO-2135.
 RA MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences."; Nucleic Acids Res. 15:9797-9805 (1987).
 RL Nucleic Acids Res. 15:9797-9805 (1987).
 RP [15]
 RX VARIANT HEMA GLN-2228.
 RA MEDLINE=88191989; PubMed=2833855;
 RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides."; Am. J. Hum. Genet. 42:718-725 (1988).
 RL Am. J. Hum. Genet. 42:718-725 (1988).
 RP [16]
 RX VARIANT HEMA GLY-291.
 RA Kazazian H.H., Antonarakis S.E.;
 RX MEDLINE=88220354; PubMed=2835904;
 RA Yousoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene."; Am. J. Hum. Genet. 42:867-871 (1988).
 RL Am. J. Hum. Genet. 42:867-871 (1988).
 RP [17]
 RX VARIANT HEMA CYS-1708.
 RA MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A."; Blood 73:2117-2122 (1989).
 RL Blood 73:2117-2122 (1989).
 RP [18]
 RX VARIANT HEMA CYS-391.
 RA MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule."; Blood 74:1612-1617 (1989).
 RL Blood 74:1612-1617 (1989).
 RP [19]
 RX VARIANT HEMA LEU-189.
 RA MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A."; Blood 74:2688-2691 (1989).
 RL Blood 74:2688-2691 (1989).
 RP [20]
 RX VARIANT HEMA LEU-2326.
 RA MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene."; Hum. Genet. 81:335-338 (1989).
 RL Hum. Genet. 81:335-338 (1989).
 RP [21]
 RX VARIANT HEMA HIS-391.
 RA MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site (arginine-
 RT 372-->histidine)"; Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281 (1989).
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281 (1989).
 RP [22]
 RX VARIANT HEMA CYS-1708.
 RA MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with cross-
 RT reacting material-positive hemophilia A."; Blood 75:384-389 (1990).
 RL Blood 75:384-389 (1990).
 RP [23]
 RX VARIANTS HEMA GLN-2228 AND LEU-2326.
 RA MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor VIII
 RT gene of hemophilia A patients of Italian descent."; Blood 75:662-670 (1990).
 RL Blood 75:662-670 (1990).
 RP [24]
 RX VARIANT HEMA CYS-391.
 RA MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site."; Br. J. Haematol. 75:73-77 (1990).
 RL Br. J. Haematol. 75:73-77 (1990).
 RP [25]
 RX VARIANTS HEMA PHE-1699 AND CYS-1708.
 RA MEDLINE=90152891; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;


```
RT "Characterization of mutations in the factor VIII gene by direct
RT sequencing of amplified genomic DNA.";

Query Match      100.0%; Score 1666; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 4.5e-132;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 60
DB 2020 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2079
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 120
DB 2080 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 2139
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 180
DB 2140 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 2199
QY 181 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKEWLQVDFOKTMKV 240
DB 2200 ESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPOVNNPKEWLQVDFOKTMKV 2259
QY 241 TGVTTQGVKSLTSMYKKEFLISSODGHOWTLFFQNGKVKVFGQNDSTFPVNSLDPP 300
DB 2260 TGVTTQGVKSLTSMYKKEFLISSODGHOWTLFFQNGKVKVFGQNDSTFPVNSLDPP 2319
QY 301 LLTRYLRHPOSW 313
DB 2320 LLTRYLRHPOSW 2332

RESULT 2
O18806 PRELIMINARY; PRT; 2343 AA.
AC O18806;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Liver;
RA Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA Giles A., Lillicrap D.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; AF016234; AAB87412.1; -.
DR HSSP; P00451; 1D7P.
DR GO; GO:0005507; F: copper ion binding; IEA.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00394; Cu-oxidase_1.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match      91.4%; Score 1522; DB 2; Length 2343;
Best Local Similarity 89.8%; Pred. No. 7.5e-120;
Matches 281; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 60
DB 2020 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2079
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 120
DB 2072 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 2131
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 180
DB 2132 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 2191

RESULT 3
O62730 PRELIMINARY; PRT; 2343 AA.
AC O62730;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Kidney;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; AF049489; AAC05384.1; -.
DR HSSP; P00451; 1IQD.
DR GO; GO:0005507; F: copper ion binding; IEA.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00394; Cu-oxidase_1.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match      91.1%; Score 1517; DB 2; Length 2343;
Best Local Similarity 89.5%; Pred. No. 2e-119;
Matches 280; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 60
DB 2012 LIGEHLAGMSTFLVYSNKQCTPLGMASGHIRDFOITASQYQGWAPKLARLHYSGSIN 2071
QY 61 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 120
DB 2072 AWSTKEPFSWIKVDLLAPMIHGIKTOGARQKFSLSYISQFIIMYSLDGKWKQTYRGNST 2131
QY 121 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 180
DB 2132 GTLMVFFGNVDSSGKHNIFNPPIIARYIRLHPTHYSIRSLRMELMCGDLNCSMPGLM 2191
```

QY 181 ESKAIDAOITASSYFTNFMATWSPSKARLHOGNSNARPOVNNPKWLQVDFQKTMKV 240
 Db 2192 ESKAIDAOITASSYFTNFMATWSPSKARLHOGNSNARPOVNNPKWLQVDFQKTMKV 2251
 QY 241 TGVTTQGVSKLITSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 300
 Db 2252 TGVTTQGVSKLITSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQDSTFTPVVNSLDPP 2311
 QY 301 LLTRYLRHPQSW 313
 Db 2312 LVARYVRLHPQSW 2324

RESULT 4

FAS_MOUSE STANDARD; PRT: 2319 AA.
 AC Q06194;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Coagulation factor VIII precursor (Procoagulant component).
 GN Name=F8; Synonyms=Cf8, F8c;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae, Murinae, Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=93300511; PubMed=8314577;
 RA Elder B., Lakich D., Gitschier J.;
 RT "Sequence of the murine factor VIII cDNA."; Genomics 16:374-379(1993).
 CC -1- FUNCTION: Factor VIII, along with calcium and phospholipid, acts as a cofactor for factor IXa when it converts factor X to the activated form, factor Xa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Found in most tissues.
 CC -1- PMW: The binding of Von Willebrand Factor (VWF) and activation depend on the sulfation of Tyr-1669.
 CC -1- SIMILARITY: Belongs to the multicopper oxidase family.
 CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L05573; AAA37385.1; -.
 CC PIR: A47004; A47004.
 CC HSP: P00451; LD7P.
 CC MGD: MGI:86383; F8.
 CC InterPro: IPR001117; Cu-oxidase.
 CC InterPro: IPR008972; Cupredoxin.
 CC InterPro: IPR000421; FA58C.
 CC Pfam: PF00394; Cu-oxidase; 1.
 CC Pfam: PF00754; F5_F8_type_C; 2.
 CC SMART: SM00231; FA58C; 2.
 CC PROSITE: PS01285; FA58C.1; 2.
 CC PROSITE: PS01286; FA58C.2; 2.
 CC PROSITE: PS00022; FA58C.3; 2.
 CC PROSITE: PS00079; MULTICOPPER OXIDASE1; 3.
 CC Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat; Signal; Sulfation.
 KW SIGNAL
 FT CHAIN 1 19 Potential.
 FT CHAIN 20 2319 Coagulation factor VIII.
 FT DOMAIN 20 349 F5/8 type A 1.
 FT DOMAIN 20 199 Plastocyanin-like 1.
 FT DOMAIN 207 349 Plastocyanin-like 2.

FT DOMAIN 399 730 F5/8 type A 2.
 FT DOMAIN 399 573 plastocyanin-like 3.
 FT DOMAIN 583 730 plastocyanin-like 4.
 B.
 FT DOMAIN 760 1640 F5/8 type A 3.
 FT DOMAIN 1683 1845 plastocyanin-like 5.
 FT DOMAIN 1855 2008 plastocyanin-like 6.
 FT DOMAIN 2008 2156 F5/8 type C 1.
 FT DOMAIN 2161 2313 F5/8 type C 2.
 FT SITE 391 352 Cleavage (by thrombin) (By similarity).
 FT SITE 759 760 Cleavage (by thrombin) (By similarity).
 FT SITE 1324 1325 Cleavage (activation) (By similarity).
 FT SITE 1640 1641 Cleavage (activation) (By similarity).
 FT SITE 1678 1679 Cleavage (by thrombin) (By similarity).
 FT MOD RES 367 367 Sulfotyrosine (By similarity).
 FT MOD RES 737 737 Sulfotyrosine (By similarity).
 FT MOD RES 738 738 Sulfotyrosine (By similarity).
 FT MOD RES 742 742 Sulfotyrosine (By similarity).
 FT MOD RES 1669 1669 Sulfotyrosine (By similarity).
 FT MOD RES 1687 1687 Sulfotyrosine (By similarity).
 FT DISULFID 173 199 Probable.
 FT DISULFID 547 573 Probable.
 FT DISULFID 1819 1845 Probable.
 FT DISULFID 2008 2156 By similarity.
 FT DISULFID 2161 2313 By similarity.
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 233 233 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 259 259 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 601 601 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 880 880 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 958 958 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1015 1015 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1022 1022 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1026 1026 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1044 1044 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1076 1076 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1087 1087 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1136 1136 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1161 1161 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1192 1192 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1255 1255 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1268 1268 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1273 1273 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1274 1274 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1302 1302 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1316 1316 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1340 1340 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1378 1378 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1797 1797 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2105 2105 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 2319 AA; 266148 MW; PD054DB051DB2A01 CRC64;
 Query Match 89.7%; Score 1495; DB 1; Length 2319;
 Best Local Similarity 89.1%; Pred. No. 1.4e-117;
 Matches 279; Conservative 13; Mismatches 21; Indels 0; Gaps 0;
 Qy 1 LIGEHLAGMSTFLVYSNKCQPTPLGMASGHIRDQITASGOYGOWAPKLARLHYSGSIN 60
 Db 1988 LIGEHLAGMSTFLVYSKQCIPLGMASGSRDQITASGYQWAPNLARLHYSGSIN 2047
 Qy 61 AWSTKEPFESWIKVDLLAPMIHGIKTQGARQKPFSSLYISQFIIMYSLDGKKWQYRGNST 120
 Db 2048 AWSTKEPFESWIKVDLLAPMIVHGKTQGARQKPFSSLYISQFIIMYSLDGKKWLSYQGNST 2107
 Qy 121 GTLMVFFGNVDSSGKINIFNPPIIARYIRLHPTHTSSIRSLRMELMGCDLNSCSPLGM 180
 Db 2108 GTLMVFFGNVDSSGKINSFNPIIARYIRLHPTHTSSIRSLRMELMGCDLNSCSPLGM 2167
 Qy 181 ESKAIDAOITASSYFTNFMATWSPSKARLHOGNSNARPOVNNPKWLQVDFQKTMKV 240
 Db 2168 ESKVLSDTQITASSYFTNFMATWSPSOARLHOGNTNARPOVNDPKWLQVLDLQKTMKV 2227

RESULT 10

Q804X4 PRELIMINARY; PRT; 745 AA.
 AC Q804X4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Coagulation factor V (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 [1]
 SEQUENCE FROM N.A.
 RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
 RA Tuddenham E.C.D., McVey J.H.,
 BL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC - SIMILARITY: Contains 2 F5/8 type C domains.
 DR EMBL; AF465271; AAC33366.1; -.
 DR HSSP; P12259; IC2T.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR00421; FA58 C.
 DR InterPro; IPR008979; Gal Bind like.
 DR SMART; SM00231; FA58C; 2.
 DR Pfam; PF00754; F5_F8 type C; 2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR PROSITE; PS50022; FA58C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
 FT NON TR 1
 SQ SEQUENCE 745 AA; 85626 MW; A264587A348C29EE CRC64;

Query Match 46.9%; Score 782; DB 2; Length 745;
 Best Local Similarity 46.8%; Pred. No. 1.2e-57;
 Matches 148; Conservative 62; Mismatches 102; Indels 4; Gaps 2;

Qy 2 IGHLLHAGMSTLFLVYNNKCTPLGMSGHIRDFQTASQYQGWAPKLARLHYSGSINA 61
 Db 412 VGEYQAGQASLYLVLEKGRCPMGASGVLDSQINASHHIDYWEKPLARLNNSGTYN 471
 Qy 62 WST--KEPFSWIKVDLLAPMIHGIKTQGARQKFSLYISQFIIMYSLDGKKWQTYRGN 119
 Db 472 WSTTTEDLPWTQVDFQVLLTGIQTQGAHFLKSFYVQKLFIVYSTDKRWTFRGDS 531
 Qy 120 TGTLMVFFGNVDSGKIHNIENPPIIARYIRLHPHYISIRSLRMELMGCDLNSCMPLG 179
 Db 532 SPAEKIFEGNSDAYGIKENIIDPPIIARYIRVYPKQAYNEPTLRMLLGCEVDACSPLG 591
 Qy 180 MESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLOVDFQKTM 239
 Db 592 MENGEIKNTQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLOVDFQKTM 239
 Qy 240 VTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGK--VKVFGNQDSFTTPVNSL 297
 Db 652 ITAIAIQGVKSTYENFVKYIVILYNSQGSWKSYTEDSSSVAKVFGNSDGRGHVKKHF 711
 Qy 298 DPPLTLRYLRHPQSW 313
 Db 712 NPPILSRFIRIVPKTW 727

RESULT 11

Q7SZNO PRELIMINARY; PRT; 1460 AA.
 AC Q7SZNO;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Pseutarin C precursor.
 OS Pseudonaja textilis (Eastern brown snake).

Query Match 46.3%; Score 771.5; DB 2; Length 1460;
 Best Local Similarity 46.7%; Pred. No. 2.1e-56;
 Matches 148; Conservative 57; Mismatches 107; Indels 5; Gaps 2;

Qy 2 IGHLLHAGMSTLFLVYNNKCTPLGMSGHIRDFQTASQYQGWAPKLARLHYSGSINA 61
 Db 1128 VGENOERGQALFTVIDKCKLPMGLASGIQDSIASGHVYWEPEKLARLNNTGKYN 1187
 Qy 62 WST--TKPFSWIKVDLLAPMIHGIKTQGARQKFSLYISQFIIMYSLDGKKWQTYRGN 118
 Db 1188 WSLIKKEHEHPWIQIDLQROVITGIGTQVQLLQHSYTFYFVTSYSEDQNNWITPKGR 1247
 Qy 119 STGTLMVFFGNVDSGKIHNIENPPIIARYIRLHPHYISIRSLRMELMGCDLNSCMPL 178
 Db 1248 HSETQHFEGNSDGTTVKENHIDPPIIARYIRLHPKTFYRNPTFRIELLGCEVGCSPVL 1307
 Qy 179 MESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKWLOVDFQKTM 238
 Db 1308 GMSGAIKNSIETASSYKKTWSSWEPSLARINLEGCTNAMOPEVNNKQDLQIDLQHLT 1367
 Qy 239 KVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTTPVNS 296
 Db 1368 KITSITQGATSMYVKEFLISSQDGHQWTLFPQNGKVKVFGNQDSFTTPVNS 296
 Qy 297 DPPLTLRYLRHPQSW 313
 Db 1428 FKPPILSRFIRIVPKTW 1444

RESULT 12

FA5_HUMAN
 ID FA5_HUMAN STANDARD; PRT; 2224 AA.
 AC P12259; Q14285;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Coagulation factor V precursor (Activated protein C cofactor).
 GN Name=F5;
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Elapidae; Acanthophiinae; Pseudonaja.
 NCBI_TaxID=8673;
 [1]
 SEQUENCE FROM N.A.
 RC TISSUE=Venom gland.
 RX MEDLINE=22781770; PubMed=12730119; DOI=10.1182/blood-2002-12-3839;
 RA Rao V.S., Svarup S., Kini R.M.;
 RA "The nonenzymatic subunit of pseutarin C, a prothrombin activator from
 eastern brown snake (Pseudonaja textilis) venom, shows structural
 similarity to mammalian coagulation factor V.";
 FT Blood 102:1347-1354(2003).
 RL - SIMILARITY: Contains 2 F5/8 type C domains.
 CC - SIMILARITY: Contains 2 F5/8 type C domains.
 DR EMBL; AY168281; AAC38805.1; -.
 DR HSSP; P12259; IC2T.
 DR MEROPS; M14_951; -.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR00421; FA58 C.
 DR InterPro; IPR008979; Gal Bind like.
 DR InterPro; IPR000834; Peptidase M14.
 DR Pfam; PF00754; F5_F8 type C; 2.
 DR SMART; SM00231; FA58C; 2.
 DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 1.
 DR PROSITE; PS50022; FA58C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_3.
 KW SIGNAL.
 FT SIGNAL 1 10 Potential.
 SQ SEQUENCE 1460 AA; 165931 MW; 6AFB63E2D5D275A6 CRC64;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT MET-1764.
RX MEDLINE=92232668; PubMed=1567832;
RA Cripe L.D., Moore K.D., Kane W.H.;
RT "Structure of the gene for human coagulation factor V.";
RL Biochemistry 31:3777-3785(1992).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT GLU-925.
RX MEDLINE=87260886; PubMed=3110773;
RA Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A.,
RW Hewick R.W., Kaufman R.J., Mann K.G.;
RT "Complete cDNA and derived amino acid sequence of human factor V.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
RN [3]
RP SEQUENCE OF 1-1600 FROM N.A., AND VARIANTS GLU-925 AND ILE-1285.
RX MEDLINE=88107560; PubMed=2827731;
RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
RT "Cloning of cDNAs coding for the heavy chain region and connecting
region of human factor V, a blood coagulation factor with four types
of internal repeats.";
RL Biochemistry 26:6508-6514(1987).
RN [4]
RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
RX MEDLINE=86313665; PubMed=3092220;
RA Kane W.H., Davie E.W.;
RT "Cloning of a cDNA coding for human factor V, a blood coagulation
factor homologous to factor VIII and ceruloplasmin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
RN [5]
RP PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=93203619; PubMed=8454869;
RA Shen N.L.L., Fan S.-T., Pyati J., Graff R., Lapolla R.J.,
RW Edgington T.S.;
RT "The serine protease cofactor factor V is synthesized by
lymphocytes.";
RL J. Immunol. 150:2992-3001(1993).
RN [6]
RP SULFATION.
RX MEDLINE=94264012; PubMed=8204629;
RA Pittman D.D., Tomkinson K.N., Michnick D., Seligsohn U.,
RW Kaufman R.J.;
RT "Posttranslational sulfation of factor V is required for efficient
thrombin cleavage and activation and for full procoagulant activity.";
RL Biochemistry 33:6952-6959(1994).
RN [7]
RP SULFATION.
RX MEDLINE=90366699; PubMed=2168225;
RA Hortin G.L.;
RT "Sulfation of tyrosine residues in coagulation factor V.";
RL Blood 76:946-952(1990).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2065-2224.
RX MEDLINE=20052169; PubMed=10586886; DOI=10.1038/46594;
RA Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M.A., Kim S.W.,
RW Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
RW Fuentes-Prior P.;
RT "Crystal structures of the membrane-binding C2 domain of human
coagulation factor V.";
RL Nature 402:434-439(1999).
RN [9]
RP VARIANT MET-1764.
RX MEDLINE=95179146; PubMed=7874144;
RA Bayston T.A., Ireland H., Olds R.J., Thein S.L., Lane D.A.;
RT "A polymorphism in the human coagulation factor V gene.";
RL Hum. Mol. Genet. 3:2085-2085(1994).
RN [10]
RP VARIANT APCR GLN-534.
RX MEDLINE=94217810; PubMed=8164741; DOI=10.1038/369064a0;
RA Bertina R.M., Koeleman B.P.C., Koster T., Rosendaal F.R., Dirven R.J.,
de Ronde H., van der Velden P.A., Reitsma P.H.;
RT "Mutation in blood coagulation factor V associated with resistance to
activated protein C.";
RL Nature 369:64-67(1994).
RN [11]
RP VARIANTS ILE-1285 AND ARG-1327.
RX MEDLINE=96351768; PubMed=8713778;
RA Lunghi B., Iacoviello L., Gemmati D., Dilasio M.G., Castoldi E.,
RA Pinotti M., Castaman G., Redaelli R., Mariani G., Marchetti G.,
RA Bernardi F.;
RT "Detection of new polymorphic markers in the factor V gene:
association with factor V levels in plasma.";
RL Thromb. Haemost. 75:45-48(1996).
RN [12]
RP VARIANT APCR THR-334.
RX MEDLINE=98122763; PubMed=9454741;
RA Chan W.P., Lee C.K., Kwong Y.L., Lam C.K., Liang R.;
RT "A novel mutation of Arg306 of factor V gene in Hong Kong Chinese.";
RL Blood 91:1135-1139(1998).
RN [13]
RP VARIANT APCR THR-334.
RX MEDLINE=98122764; PubMed=9454742;
RA Williamson D., Brown K., Luddington R., Baglin C., Baglin T.;
RT "Factor V Cambridge: a new mutation (Arg306-to-Thr) associated with
resistance to activated protein C.";
RL Blood 91:1140-1144(1998).
RN [14]
RP VARIANTS HIS-107; THR-413; LYS-513; SER-809; THR-817; ARG-858;
RX ARG-865; GLN-925; GLN-1146; ALA-1530; SER-1685; VAL-1749; MET-1764;
RP ILE-1820 AND GLY-2222, AND VARIANT APCR GLN-534;
RX MEDLINE=99318093; PubMed=10391209; DOI=10.1038/10290;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RW Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [15]
RP ERRATUM.
RX PubMed=10545957;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RW Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
RN [16]
RP VARIANT APCR HIS-2102.
RX MEDLINE=21847288; PubMed=11858490;
RA Schrijver I., Houissa-Kastally R., Jones C.D., Garcia K.C.,
RW Zehnder J.L.;
RT "Novel factor V C2-domain mutation (R2074H) in two families with
factor V deficiency and bleeding.";
RL Thromb. Haemost. 87:294-299(2002).
RN CC -|- FUNCTION: Coagulation factor V is a cofactor that participates
CC with factor Xa to activate prothrombin to thrombin.
CC CC -|- SUBUNIT: Factor Va is composed of a heavy chain and a light chain,
CC noncovalently bound. The interaction between the two chains is
CC calcium-dependent.
CC CC -|- DOMAIN: Domain B contains 35 x 9 AA tandem repeats, and 2 x 17 AA
CC repeats.
CC CC -|- PTM: Thrombin activates factor V proteolytically to the active
CC cofactor, factor Va (formation of a heavy chain at the N-terminus
CC and a light chain at the C-terminus).
CC CC -|- PTM: Sulfation is required for efficient thrombin cleavage and
CC activation and for full procoagulant activity.
CC CC -|- DISEASE: Defects in F5 are the cause of Owren parahemophilia
CC [MIM:227400], an hemorrhagic diathesis.
CC CC -|- DISEASE: Defects in F5 are the cause of resistance to activated
CC protein C (APCR) [MIM:188055], a form of thrombophilia. The APCR
CC mutation is found in about 5% of the population which suggest that
CC a slight thrombotic tendency may confer some advantage in fetal
CC implantation.

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN Name=F5;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING OF F5/8 TYPE A AND C
RP DOMAINS.
RC TISSUE=Liver;
RX MEDLINE=21121490; PubMed=11229814;
RA Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
RA Kim H.K.W.;
RT "Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT protein modeling of membrane binding sites and comparative anatomy of
RT domains.";
RL Cell. Mol. Life Sci. 58:148-159(2001).
CC -!- FUNCTION: Coagulation factor V is a cofactor that participates
CC with factor Xa to activate prothrombin to thrombin.
CC -!- SUBUNIT: Factor Va is composed of a heavy chain and a light chain,
CC noncovalently bound. The interaction between the two chains is
CC calcium-dependent.
CC -!- DOMAIN: Domain B contains 41 X 9 AA tandem repeats. Domains C1 and
CC C2 may be involved in membrane binding.
CC -!- PTM: Thrombin activates factor V proteolytically to the active
CC cofactor, factor Va (formation of a heavy chain at the N-terminus
CC and a light chain at the C-terminus).
CC -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF191308; AAC28381.1; --
DR HSPF; P12259; ICZT.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR008979; Gal Bind_like.
DR InterPro; IPR009271; LSPR; 40.
DR Pfam; PF00754; F5_F8 type_C; 2.
DR Pfam; PF06049; LSPR; 40.
DR PROSITE; PS01285; FA58C 1; 2.
DR PROSITE; PS01286; FA58C 2; 2.
DR PROSITE; PS00022; FA58C 3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
KW Blood coagulation; Calcium; Glycoprotein; Repeat; Signal; Sulfation;
KW Zymogen.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 2258 Coagulation factor V.
FT CHAIN 23 737 Coagulation factor V heavy chain (By
FT similarity).
FT PROPEP 738 1611 Activation peptide (connecting region)
FT (By similarity).
FT CHAIN 1612 2258 Coagulation factor V light chain (By
FT similarity).
FT DOMAIN 30 329 F5/8 type A 1.
FT DOMAIN 30 193 Plastocyanin-like 1.
FT DOMAIN 203 329 Plastocyanin-like 2.
FT DOMAIN 348 683 F5/8 type A 2.
FT DOMAIN 348 525 Plastocyanin-like 3.
FT DOMAIN 535 683 Plastocyanin-like 4.
FT DOMAIN 691 1611 B.
FT DOMAIN 1168 1539 41 X 9 AA approximate tandem repeats of
FT T-L-S-P-D-L-[GS]-[HQ]-T.

FT REPEAT 1168 1176 1.
FT REPEAT 1177 1185 2.
FT REPEAT 1186 1194 3.
FT REPEAT 1195 1203 4.
FT REPEAT 1204 1212 5.
FT REPEAT 1213 1221 6.
FT REPEAT 1222 1230 7.
FT REPEAT 1231 1239 8.
FT REPEAT 1240 1248 9.
FT REPEAT 1249 1257 10.
FT REPEAT 1258 1266 11.
FT REPEAT 1267 1275 12.
FT REPEAT 1276 1284 13.
FT REPEAT 1285 1293 14.
FT REPEAT 1294 1302 15.
FT REPEAT 1303 1311 16.
FT REPEAT 1312 1320 17.
FT REPEAT 1321 1329 18.
FT REPEAT 1330 1338 19.
FT REPEAT 1339 1347 20.
FT REPEAT 1348 1356 21.
FT REPEAT 1357 1365 22.
FT REPEAT 1366 1374 23.
FT REPEAT 1375 1383 24.
FT REPEAT 1384 1392 25.
FT REPEAT 1393 1401 26.
FT REPEAT 1402 1410 27.
FT REPEAT 1411 1419 28.
FT REPEAT 1420 1428 29.
FT REPEAT 1429 1437 30.
FT REPEAT 1438 1446 31.
FT REPEAT 1447 1455 32.
FT REPEAT 1456 1464 33.
FT REPEAT 1465 1473 34.
FT REPEAT 1474 1482 35.
FT REPEAT 1483 1491 36.
FT REPEAT 1492 1500 37.
FT REPEAT 1501 1509 38.
FT REPEAT 1510 1518 39.
FT REPEAT 1519 1527 40.
FT REPEAT 1531 1539 41.
FT DOMAIN 1616 1941 F5/8 type A 3.
FT DOMAIN 1616 1785 Plastocyanin-like 5.
FT DOMAIN 1795 1941 Plastocyanin-like 6.
FT DOMAIN 1941 2095 F5/8 type C 1.
FT DOMAIN 2100 2255 F5/8 type C 2.
FT SITE 737 738 Cleavage (by thrombin) (By similarity).
FT SITE 1029 1030 Cleavage (by thrombin) (By similarity).
FT SITE 1611 1612 Cleavage (by thrombin) (By similarity).
FT DISULFID 167 193 By similarity.
FT DISULFID 248 329 By similarity.
FT DISULFID 499 525 By similarity.
FT DISULFID 602 683 By similarity.
FT DISULFID 1759 1785 Probable.
FT DISULFID 1941 2095 By similarity.
FT DISULFID 2100 2255 By similarity.
FT MOD_RES 692 692 Sulfotyrosine (Potential).
FT MOD_RES 696 696 Sulfotyrosine (Potential).
FT MOD_RES 724 724 Sulfotyrosine (Potential).
FT MOD_RES 726 726 Sulfotyrosine (Potential).
FT MOD_RES 745 745 Sulfotyrosine (Potential).
FT MOD_RES 1560 1560 Sulfotyrosine (Potential).
FT MOD_RES 1576 1576 Sulfotyrosine (Potential).
FT MOD_RES 1581 1581 Sulfotyrosine (Potential).
FT MOD_RES 1584 1584 Sulfotyrosine (Potential).
FT MOD_RES 1588 1588 Sulfotyrosine (Potential).
FT MOD_RES 1631 1631 Sulfotyrosine (Potential).
FT CARBOHYD 153 153 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 225 225 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 239 239 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 297 297 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 382 382 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 460 460 N-linked (GlcNAc...) (Potential).

```

FT CARBOHYD 467 467 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 553 553 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 741 741 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 752 752 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 760 760 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 776 776 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 782 782 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 899 899 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 960 960 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1048 1048 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1057 1057 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1066 1066 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1073 1073 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1089 1089 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1174 1174 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1480 1480 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1537 1537 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1597 1597 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1737 1737 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 1886 1886 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 2044 2044 N-linked (GlcNAc. .) (potential).
FT CARBOHYD 2243 2243 N-linked (GlcNAc. .) (potential).
SQ SEQUENCE 2258 AA; 256078 MW; 9159B9E0076A2ACC CRC64;

Query Match 44.6%; Score 743.5; DB 1; Length 2259;
Best Local Similarity 44.4%; Pred. No. 8.4e-54;
Matches 143; Conservative 63; Mismatches 105; Indels 11; Gaps 4;

QY 2 IGEHLHAGMSTLFLVYSNKCQTPLGMAHGHRDQITASQYQGWAPKLARLHYSGSINA 61
DB 1922 VGENQVAGMOTPELLIDRECKMPGLSTGLIADSQLKASEFHWQPKLARLNNGGSYNA 1981
QY 62 WSTKEPFS-----WIKVDLLAPMIHGIKTQAGOKFSSLYISQFIIMYSLDGGKKWOT 114
DB 1982 WIT-DKFSGESNKEPWQVDMQREVVFTGQTQAKYLYKSYVTFEFNAVYSSDORNRI 2040
QY 115 YRGNSTGTLVFGNVDSGKHNIFNPPIIARYIRLHPHYSIRSTRMELMGCGLNSC 174
DB 2041 FKNSTKNWYFNNGSDASTITENQDPFPPVARYIRISPTESYNKPALRLLELQCEVNGC 2100
QY 175 SMLPGHESKAISDAQITASSYFTNWAT-WSPSKARLHLOGRSNARWQVNPKNKWLQVD 233
DB 2101 STPLGMEGNKNEQITASSFKKSWGDYWEPPRLNAGQVRNANQAKANNNNQWLQID 2160
QY 234 FQTKMKTGVTGQVKSLTSMYKFEFLISSQDGHQWTLFFQNGKV--KVFQGNQDSFT 291
DB 2161 LKIKKITAIVTQCKSLSSSEMYVRRITQYSDRGVENSREKSSMVDKIFEGNNGK 2220
QY 292 PVNSLDPLLRILRIHPQSW 313
DB 2221 HVKNFNPPIISFRIRIIPKMW 2242

```

RESULT 15

```

O88783 ID O88783 PRELIMINARY; PRT; 2183 AA.
AC O88783;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Coagulation factor V.
GN Name=F5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=98282202; PubMed=9616155;
RA Yang T.L., Cui J., Rehmtulla A., Yang A., Moussalli M., Kaufman R.J.,
RA Ginsburg D.;
RT "The structure and function of murine factor V and its inactivation by
RT protein C.";

```

```

RL Blood 91:4593-4599(1998).
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR EMBL; U52925; AAC99553.1; -.
DR PIR; T42764; T42764.
DR HSSP; P12259; 1CZT.
DR MGD; MGI-88382; F5.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001117; Cu oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000431; FAS8_C.
DR InterPro; IPR009271; LSPD_C.
DR Pfam; PF00754; F5_F8_Type_C; 2.
DR SMART; SM00231; FAS8C; 2.
DR PROSITE; PS01285; FAS8C_1; 2.
DR PROSITE; PS01286; FAS8C_2; 2.
DR PROSITE; PS00022; FAS8C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ SEQUENCE 2183 AA; 247228 MW; BF0A8AA723F60317 CRC64;

Query Match 44.5%; Score 741.5; DB 2; Length 2183;
Best Local Similarity 45.2%; Pred. No. 1.12e-53;
Matches 145; Conservative 60; Mismatches 107; Indels 9; Gaps 4;

QY 2 IGEHLHAGMSTLFLVYSNKCQTPLGMAHGHRDQITASQYQGWAPKLARLHYSGSINA 61
DB 1847 VGENQVAGMOTPELLIDRECKMPGLSTGLIADSQLKASEYLTWPELRLNNGSYNA 1906
QY 62 WSTKEPFS-----WIKVDLLAPMIHGIKTQAGOKFSSLYISQFIIMYSLDGGKKWOT 115
DB 1907 WSIKLTALDFPIKPIQVDMQREVVFTGQTQAKYLYKSCFTTEFQVAYSSDQTNQIF 1966
QY 116 RGNSTGTLVFGNVDSGKHNIFNPPIIARYIRLHPHYSIRSTRMELMGCGLNSC 175
DB 1967 RKGSGKSWYFTGNSDGTIKENRLDPPVARYIRIHPKSYNRPFLRLLELQCEVNGCS 2026
QY 176 SMLPGHESKAISDAQITASSYFTNWAT-WSPSKARLHLOGRSNARWQVNPKNKWLQVD 234
DB 2027 TPLGLEDGRIDQKITASSFKKSWGDYWEPPRLNAGQVRNANQAKANNNNQWLQID 2086
QY 235 QTKMKTGVTGQVKSLTSMYKFEFLISSQDGHQWTLFFQNGKV--KVFQGNQDSFT 292
DB 2087 LKIKKITAIVTQCKSLSSSEMYVRRITQYSDRGVANSYKSIQYSDQGVANWPKPYRQKSSMVDKIFEGNSNTKG 2146
QY 293 VVNSLDPLLRILRIHPQSW 313
DB 2147 MKNFNPPIISFRIRIIPKMW 2167

Search completed: April 20, 2005, 17:21:03
Job time : 80.25 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:08:58 ; Search time 28.3812 Seconds
(without alignments)
1240.089 Million cell updates/sec

Title: US-09-853-080b-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPEKWLQVDQFKTMKVTG.....VNSLDPLLTLYLRHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	100.0	160	2 AAW69527	Aaw69527 C2 peptid
2	484	100.0	216	8 ADQ37592	Adq37592 Human myo
3	484	100.0	642	6 ABP60515	Abp60515 Human fac
4	484	100.0	643	6 ABP60514	Abp60514 Human fac
5	484	100.0	684	2 AAR73022	Aar73022 Human fac
6	484	100.0	684	2 AAR74091	Aar74091 Factor-VI
7	484	100.0	770	3 ADP20767	Adp20767 Factor VI
8	484	100.0	790	5 ADE64594	Ad64594 Recombina
9	484	100.0	1014	1 AAP71139	Aap71139 Factor VII
10	484	100.0	1383	2 AAW33227	Aaw33227 Procoagul
11	484	100.0	1383	2 AAW33228	Aaw33228 Procoagul
12	484	100.0	1383	2 AAW33229	Aaw33229 Procoagul
13	484	100.0	1421	8 ADQ37613	Adq37613 Human fac
14	484	100.0	1421	8 ADQ37598	Adq37598 Human fac
15	484	100.0	1422	8 ADQ37591	Adq37591 Human fac
16	484	100.0	1422	8 ADQ37606	Adq37606 Human fac
17	484	100.0	1424	1 AAP80268	Aap80268 Modified
18	484	100.0	1424	1 AAP91169	Aap91169 Sequence
19	484	100.0	1424	4 AAB48842	Aab48842 Mutant ma
20	484	100.0	1424	5 AAC18622	Aac18622 Human mat
21	484	100.0	1425	1 AAP80267	Aap80267 Modified
22	484	100.0	1425	8 ADQ37605	Adq37605 Human fac
23	484	100.0	1425	8 ADQ37590	Adq37590 Human fac
24	484	100.0	1428	8 ADQ37599	Adq37599 Human fac
25	484	100.0	1428	8 ADQ37604	Adq37604 Human fac

26	484	100.0	1428	8 ADQ37589	Adq37589 Human fac
27	484	100.0	1428	8 ADQ37614	Adq37614 Human fac
28	484	100.0	1434	8 ADQ37594	Adq37594 Human fac
29	484	100.0	1434	8 ADQ37609	Adq37609 Human fac
30	484	100.0	1435	8 ADQ37600	Adq37600 Human fac
31	484	100.0	1435	8 ADQ37615	Adq37615 Human fac
32	484	100.0	1437	8 ADQ37593	Adq37593 Human fac
33	484	100.0	1437	8 ADQ37608	Adq37608 Human fac
34	484	100.0	1438	3 AAB01262	Aab01262 B-domain
35	484	100.0	1438	7 ADF31316	Adf31316 Variant o
36	484	100.0	1438	7 ADM75988	Adm75988 Modified
37	484	100.0	1438	7 ADM75985	Adm75985 Modified
38	484	100.0	1438	7 ADM75986	Adm75986 Modified
39	484	100.0	1438	7 ADM75989	Adm75989 Modified
40	484	100.0	1438	7 ADM75983	Adm75983 Wild-type
41	484	100.0	1438	7 ADM75984	Adm75984 Modified
42	484	100.0	1438	7 ADM75987	Adm75987 Modified
43	484	100.0	1440	2 AAR12971	Aar12971 Factor VI
44	484	100.0	1440	8 ADQ37607	Adq37607 Human fac
45	484	100.0	1440	8 ADQ37592	Adq37592 Human fac

ALIGNMENTS

RESULT 1
AAW69527
ID AAW69527 standard; peptide; 160 AA.
XX
AC AAW69527;
XX
DT 09-NOV-1998 (first entry)
XX
DE C2 peptide SEQ ID NO:9 from WO9829453 Example 3.
XX
KW Drug; cell membrane-directed drug; phospholipid; lipid bilayer;
KW cell cortex; blood coagulation; inflammation; immunological disorder.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9829453-A1.
XX
PD 09-JUL-1998.
XX
PF 05-JAN-1998; 98WO-JP0000002.
XX
PR 27-DEC-1996; 96JP-00359053.
XX
PA (MOCH) MOCHIDA PHARM CO LTD.
XX
PI Kuriyama S, Hasegawa T;
XX
DR WPI; 1998-388051/33.
DR N-ESDB; AAV40051.
XX
PT Drugs containing peptide(s) with specific affinity to phospholipid(s) -
PT such as phosphatidyl serine, for treatment of blood coagulation,
PT inflammatory and immunological disorders.
XX
PS Example 3; Page 72-73; 117pp; Japanese.
XX
CC The present invention describes drug compositions which contain as an
CC active component a peptide which has specific affinity to particular
CC phospholipids (such as phosphatidyl ethanolamine or phosphatidyl serine),
CC especially to phospholipids which constitute a lipid bilayer of cellular
CC cortex and of which the concentration in the bilayer increases in cells
CC which are abnormal (e.g. through injury, denaturation or activation). In
CC particular, the peptide contains a sequence having phospholipid affinity
CC and a structure of formula (I): (Al)a-(A2)b-(A3)c, where (Al) is one of
CC two specific sequences (see AAW69516 and AAW69519), (A2) and (A3) are
CC TRYLRHPQSWVHQIALR, LRLRLHPQSWVHQIALR (see AAW69517) or MEVLGCEAQNLY
CC (see AAW69518); a = 0-5; b = 1-5, and c = 0-5. Preferred are the formulae

CC A1-A2-A3, A2-A3, A2-A2-A3, A2-A2-A3 or A2-A2 (especially A2-A2-A3, A2-A2-A3 or A2-A2). The sequence is linked to a peptide such as a blood factor, especially Thrombo-modulin, urina-statin or membrane cofactor protein. The drugs are used for the treatment and prevention of diseases involving blood coagulation, inflammatory and immunological disorders. CC The present sequence represents a peptide from the present invention XX

XX Sequence 160 AA;

Query Match 100.0%; Score 484; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 9.5e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTVGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 60
DB 51 VNNPKWLQVDFQKTMKVTVGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 110

QY 61 FQGNQDSFTPVVNSLDPLLRILYLRHPQSW 91
DB 111 FQGNQDSFTPVVNSLDPLLRILYLRHPQSW 141

RESULT 2
ADQ39370 standard; protein; 216 AA.

XX ADQ39370;
XX AC
XX DT 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1033.
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX cardiact; gene therapy; human.

XX Homo sapiens.
XX WO2004058052-A2.
XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.
XX 20-DEC-2002; 2003US-0434778P.
XX 10-MAR-2003; 2003US-0453135P.
XX 30-APR-2003; 2003US-0466412P.
XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX N-PSDB; ADQ38542.

XX Identifying an individual who has an altered risk for developing
XX myocardial infarction by detecting a single nucleotide polymorphism in
XX the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1033; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
XX has an altered risk for developing myocardial infarction. The method
XX comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX the nucleotide sequences given in the specification in the individual's
XX nucleic acids, where the presence of the SNP is correlated with an
XX altered risk for myocardial infarction in the individual. The invention
XX further comprises: an isolated nucleic acid molecule comprising at least
XX 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX the specification or its complement and encoding any one of the amino
XX acid sequences given in the specification; an isolated polypeptide
XX comprising an amino acid sequence given in the specification; an antibody
XX that specifically binds to the polypeptide or its antigen-binding

CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIFO website.

XX Sequence 216 AA;

Query Match 100.0%; Score 484; DB 8; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.4e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTVGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 60
DB 107 VNNPKWLQVDFQKTMKVTVGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 166

QY 61 FQGNQDSFTPVVNSLDPLLRILYLRHPQSW 91
DB 167 FQGNQDSFTPVVNSLDPLLRILYLRHPQSW 197

RESULT 3
ABP60515 standard; protein; 642 AA.

XX ABP60515;
XX AC
XX DT 28-MAR-2003 (first entry)
XX Human factor VIII light chain A2201 deletion mutant.

XX Human; factor VIII; T-cell; haemostatic; gene therapy;
XX thrombotic disorder; haemophilia; mutant; mutein.

XX Homo sapiens.
XX WO200298454-A2.
XX 12-DEC-2002.
XX 31-MAY-2002; 2002WO-BE000088.
XX 31-MAY-2001; 2001WO-EP006297.
XX 14-NOV-2001; 2001GB-00027321.

XX (COLL-) COLLEN RES FOUND VZW ORDERWIJSEN NAVORSI.
XX Jacquemin MG, Saint-Remy JR;
XX WPI; 2003-140553/13.
XX N-PSDB; ABV99882.

XX New recombinant modified functional polypeptide, useful as a medicament
XX for treating or preventing hemophilia A or B, comprises multiple point-
XX mutations responsible for T-cell activation reduction or elimination.
XX Claim 90; Page 78-81; 85pp; English.

XX The invention relates to a novel recombinant modified functional
XX polypeptide which exerts at least one function of a mammalian protein and
XX which eliminates or reduces by at least about 80% with respect to
XX activation by the unmodified polypeptide, the activation of at least one
XX T-cell clone derived from a mammal with antibody against the unmodified

CC polypeptide. The polypeptide of the invention has haemostatic activity.
 CC The polynucleotide encoding the polypeptide may have a use in gene
 CC therapy. The recombinant modified functional polypeptide is useful as a
 CC medicament for treating or preventing a thrombotic disorder, such as
 CC haemophilia A or B. The peptide is also useful for in vitro or in vivo
 CC evaluation and/or modulation of T-cell reactivity. The recombinant point-
 CC mutated factor VIII light chain molecule is used as a template for
 CC determining whether modification of the factor VIII molecule is able to
 CC reduce the clearance, e.g. to increase the half-time, of factor VIII in
 CC plasma. The present sequence represents the human factor VIII light chain
 CC A2201 deletion mutant
 XX
 SQ Sequence 642 AA;

Query Match 100.0%; Score 484; DB 6; Length 642;
 Best Local Similarity 100.0%; Pred. No. 6.2e-51;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
 Db 533 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 592
 Qy 61 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 91
 Db 593 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 623

RESULT 4
 ABP60514
 ID ABP60514 standard; protein; 643 AA.
 XX
 AC ABP60514;
 XX
 DT 28-MAR-2003 (first entry)
 DE Human factor VIII light chain P2153Q mutant.
 XX
 KW Human; factor VIII; T-cell; haemostatic; gene therapy; light chain;
 KW thrombotic disorder; haemophilia; mutant; mutuin.
 XX
 OS Homo sapiens.
 XX
 PN WO200298454-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-BE000088.
 XX
 PR 31-MAY-2001; 2001WO-EP006297.
 PR 14-NOV-2001; 2001GB-00027321.
 XX
 PA (COLL-) COLLEN RES FOUND VZW ONDERWIJSEN NAVORSI.

XX Jacquemin MG, Saint-Remy JR;
 XX WPI; 2003-140553/13.
 XX N-PSDB; ABV99881.
 XX
 PT New recombinant modified functional polypeptide, useful as a medicament
 PT for treating or preventing hemophilia A or B, comprises multiple point-
 PT mutations responsible for T-cell activation reduction or elimination.
 XX
 PS Claim 79; Page 73-75; 85pp; English.
 XX
 CC The invention relates to a novel recombinant modified functional
 CC polypeptide which exerts at least one function of a mammalian protein and
 CC which eliminates or reduces by at least about 80%, with respect to
 CC activation by the unmodified polypeptide, the activation of at least one
 CC T-cell clone derived from a mammal with antibody against the unmodified
 CC polypeptide. The polypeptide of the invention has haemostatic activity.
 CC The polynucleotide encoding the polypeptide may have a use in gene
 CC therapy. The recombinant modified functional polypeptide is useful as a
 CC medicament for treating or preventing a thrombotic disorder, such as

CC haemophilia A or B. The peptide is also useful for in vitro or in vivo
 CC evaluation and/or modulation of T-cell reactivity. The recombinant point-
 CC mutated factor VIII light chain molecule is used as a template for
 CC determining whether modification of the factor VIII molecule is able to
 CC reduce the clearance, e.g. to increase the half-time, of factor VIII in
 CC plasma. The present sequence represents the human factor VIII light chain
 CC P2153Q mutant
 XX
 SQ Sequence 643 AA;

Query Match 100.0%; Score 484; DB 6; Length 643;
 Best Local Similarity 100.0%; Pred. No. 6.3e-51;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
 Db 534 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 593
 Qy 61 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 91
 Db 594 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 624

RESULT 5
 AAR73022
 ID AAR73022 standard; peptide; 684 AA.
 XX
 AC AAR73022;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-NOV-1995 (first entry)
 XX
 DE Human Factor-VIII C-terminal fragment.
 XX
 KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
 XX
 OS Homo sapiens.
 XX
 PN WO9513301-A1.
 XX
 PD 18-MAY-1995.
 XX
 PF 10-NOV-1994; 94WO-DK000424.
 XX
 PR 12-NOV-1993; 93DK-00001281.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Persson E;
 XX
 DR WPI; 1995-194038/25.
 XX
 PT Crosslinked Factor VIII polypeptide which is stable - is prepd. using
 PT bis(sulphosuccinimidylyl) suberate or disuccinimidylyl suberate in the
 PT presence of polysorbate 80 to produce a coagulant with long lasting
 PT activity.
 XX
 PS Disclosure; Page 24; 36pp; English.
 XX
 CC This is the C-terminal fragment (corresponding to AAs 1649-2332) of human
 CC Factor-VIII which may be crosslinked resulting in increased stability and
 CC retention of high activity over extended periods of time after activation
 CC by thrombin. The polypeptide is used to prevent or treat diseases caused
 CC by the absence or deficiency of Factor-VIII in a subject such as
 CC haemophilia. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 684 AA;

Query Match 100.0%; Score 484; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 6.8e-51;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60

Db 575 VNNPEWLQVDFQTKMKVGTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 634
Qy 61 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 91
Db 635 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 665

RESULT 6

AAR74091
ID AAR74091 standard; protein; 684 AA.
AC AAR74091;
XX
XX
DT 25-MAR-2003 (revised)
DT 04-NOV-1995 (first entry)
DE Factor-VIII light chain C-terminal fragment.
XX human; Factor VIII; light chain; C-terminal fragment; thrombin cleavage;
KW blood-clotting.
XX Homo sapiens.
OS
XX WO9513300-A1.
PN
XX 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK000423.
XX
PR 12-NOV-1993; 93DK-00001280.
XX
XX (NOVO) NOVO-NORDISK AS.
PA
XX
XX Kjalke M, Ezban Rasmussen M;
PI WPI; 1995-194037/25.
DR
XX
XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
PT activity, may be used to treat patients who have developed antibodies to
PT C-terminal epitope(s) of Factor VIII.
XX
XX Claim 4; Page 33-35; 51pp; English.

XX The sequence represents C-terminal residues 1649-2332 of a human Factor-
XX VIII light chain. The sequence is produced by treating a full-length
XX Factor-VIII polypeptide with a protease, e.g. thrombin. The fragment may
XX be produced recombinantly in conjunction with a C-terminally truncated
XX heavy chain to reduce production costs and improve safety, and production
XX levels and stability are higher than for the full-length form. When the
XX fragment is used along with a C-terminally truncated heavy chain, it may
XX be used to treat patients who have developed antibodies against epitopes
XX in the C-terminal part of the heavy chain. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX

Query Match

XX Sequence 684 AA;
XX
XX Query Match 100.0%; Score 484; DB 2; Length 684;
XX Best Local Similarity 100.0%; Pred. No. 6.8e-51;
XX Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPEWLQVDFQTKMKVGTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
Db 575 VNNPEWLQVDFQTKMKVGTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 634
Qy 61 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 91
Db 635 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 665

RESULT 7

ADP20767

ID ADP20767 standard; protein; 770 AA.
XX
AC ADP20767;
XX
DT 12-AUG-2004 (first entry)
XX
DE Factor VIII protein.
XX
XX A-domain; C-domain; blood coagulation; factor VIII;
KW von Willebrand factor.
XX
OS Unidentified.
XX
XX KR99066381-A.
PN
XX 16-AUG-1999.
PD
XX 24-JAN-1998; 98KR-00002264.
PF
XX 24-JAN-1998; 98KR-00002264.
PR
XX (GREG) KOREA GREEN CROSS CORP.
PA
XX Kim HC, Byun TH, Kim SM, Kim JM, Lee SG;
PI
XX WPI; 2000-547435/50.
DR N-PSDB; ADP20766.
XX
XX Cell line expressing A-domain, C-domain of blood coagulation factor VIII
PT and von Willebrand factor simultaneously, method - for manufacturing
PT recombinant blood coagulation factor VIII using the same.
XX
XX Disclosure; Fig 2; 21pp; Korean.
XX
XX The invention relates to a novel cell line expressing A-domain, C-domain
CC of blood coagulation factor VIII and von Willebrand factor
CC simultaneously, and a method for manufacturing recombinant blood
CC coagulation factor VIII using the same. The present sequence represents a
XX factor VIII protein.
XX
XX Sequence 770 AA;

Query Match 100.0%; Score 484; DB 3; Length 770;
Best Local Similarity 100.0%; Pred. No. 8e-51; 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 0; Indels 0;

Qy 1 VNNPEWLQVDFQTKMKVGTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
Db 661 VNNPEWLQVDFQTKMKVGTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKVKV 720
Qy 61 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 91
Db 721 FQGNQDSFTPVVNSLDPPLTLRYLRIHPQSW 751

RESULT 8

ADE64594
ID ADE64594 standard; protein; 790 AA.
XX
AC ADE64594;
XX

29-JAN-2004 (first entry)

XX Recombinant blood coagulation factor VIII protein, SEQ ID 4.
XX blood coagulation factor VIII; type-A haemophilia.
XX Unidentified.
XX
XX CN1361178-A.
XX
XX 31-JUL-2002.
XX

PF 29-DEC-2000; 2000CN-00137779.
XX
PR 29-DEC-2000; 2000CN-00137779.
XX
PA (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
XX
PI Qi Z, Wang Q, Chen C;
XX
XX WPI; 2002-741852/81.
DR N-PSDB; ADE64593.
XX
PT New recombinant blood coagulation factor VIII and its production process
XX and medicinal composition.
PT
PS Claim 2; SEQ ID NO 4; 31pp; Chinese.
XX
XX The invention relates to a novel recombinant blood coagulation factor
CC VIII, its production process and its medicinal composite for treating
CC type-A haemophilia. The invention further comprises a medicinal
CC composition containing the blood coagulation factor which promotes blood
CC coagulation to the blood plasma of type-A haemophilia patients. This
CC sequence represents a recombinant blood coagulation factor VIII protein
CC of the invention.
XX
SQ Sequence 790 AA;
Query Match 100.0%; Score 484; DB 5; Length 790;
Best Local Similarity 100.0%; Pred. No. 8.3e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
DB 681 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 740
QY 61 FQGNQDSFTPVNSLDPPLTRYLRIHPQSW 91
DB 741 FQGNQDSFTPVNSLDPPLTRYLRIHPQSW 771
RESULT 9
AAP71139
ID AAP71139 standard; protein; 1014 AA.
XX
AC AAP71139;
XX
DT 25-MAR-2003 (revised)
DT 14-MAY-1991 (first entry)
XX
DE Factor VIII:c variant.
XX
XX Factor VIII:c; variant; proteolysis; resistance;
KW pro coagulation activity.
XX
OS Homo sapiens.
XX
PN W0807144-A.
XX
PD 03-DEC-1987.
XX
PF 29-MAY-1987; 87WO-US001299.
XX
PR 29-MAY-1986; 86US-00868410.
PR 18-NOV-1986; 86US-00932767.
PR 09-DEC-1986; 86US-00939658.
XX
PA (GEMY) GENETICS INST INC.
XX
XX Kaufman RJ, Pittman D, Toole JJ;
PI WPI; 1987-348539/49.
XX
DR WPI; 1987-348539/49.
XX
PT New deletion and replacement variants of factor 8:c - resistant to
PT proteolysis but retaining pro-coagulant activity, and new DNA coding

PT sequences.
XX Disclosure; Page 1; 42pp; English.
XX
CC The full-length human factor VIII:c cDNA has been set forth in detail in
CC WO8501961. This sequence is an example of the formula A-X-B, wherein
CC A=Ala1-Arg372, B=Ser1690-Tyr2332 and X=0-1316 amino acids substantially
CC duplicative of sequences of amino acids within the sequence Arg372-
CC Ser1690 of the full-length sequence. Here X=0; producing a Arg371-Ser1690
CC fusion protein. One or more deletions or replacements at Arg 220, 226,
CC 279, 282, 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
CC variants which are more resistant to specific proteolytic cleavage
CC compared with natural factor VIII:c. Pro-coagulant activity and thrombin
CC activatability have been retained. See also AAP71726-29. (Updated on 25-
CC MAR-2003 to correct PA field.)
XX
SQ Sequence 1014 AA;
Query Match 100.0%; Score 484; DB 1; Length 1014;
Best Local Similarity 100.0%; Pred. No. 1.2e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
DB 905 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 964
QY 61 FQGNQDSFTPVNSLDPPLTRYLRIHPQSW 91
DB 965 FQGNQDSFTPVNSLDPPLTRYLRIHPQSW 995
RESULT 10
AAW33227
ID AAW33227 standard; protein; 1383 AA.
XX
AC AAW33227;
XX
DT 21-OCT-2004 (revised)
DT 30-APR-1998 (first entry)
XX
DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..346 /note= "factor VIIIA heavy chain"
FT Domain /note= "A1 domain"
FT Domain /note= "A1 domain"
FT Domain /note= "plastocyanin-like domain 1"
FT Disulfide-bond 153..179 /note= "probable"
FT Domain 187..329 /note= "plastocyanin-like domain 2"
FT Cleavage-site 372..373 /note= "by thrombin"
FT Domain 380..711 /note= "A2 domain"
FT Domain 380..554 /note= "plastocyanin-like domain 3"
FT Disulfide-bond 528..554 /note= "probable"
FT Domain 564..711 /note= "plastocyanin-like domain 4"
FT Region 711..746

XX Kaufman RJ, Pipe SW, Amano K;
 XX WPI; 1997-535830/49.
 XX Modified human pro-coagulant active factor VIII - can be administered to
 XX haemophiliacs, i.e. factor VIII replacement therapy.
 XX Claim 18; Page; 57pp; English.
 XX The present sequence represents a novel pro-coagulant active factor VIII
 CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 CC Willebrand factor binding site, mutations R336I, R562K and R740A and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains.
 CC Factor VIII, along with calcium and phospholipid, acts as a cofactor for
 CC factor IXa, when it converts factor X to the activated form (factor Xa).
 CC FVIII is the coagulation factor deficient in the X-chromosome-linked
 CC bleeding disorder haemophilia A. Several other mutant FVIII proteins have
 CC also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
 CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII mutant
 CC R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein
 CC C (APC) cleavage. The FVIII mutant comprising a deletion of the B domain
 CC and von Willebrand factor binding site, a mutation at Arg740 and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains
 CC can form a more stable configuration, and have an approximate 5-fold
 CC increase in specific activity compared to purified wild type FVIII, while
 CC increasing their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy. note: this sequence does not appear in the
 CC specification; it was created using sequences from the given references
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to feature table key
 XX Sequence 1383 AA;
 SQ
 Query Match 100.0%; Score 484; DB 2; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 1.8e-50;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VNNPEWLQVDFQTKMTKVTGTTQGVKSLLSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
 DB 1274 VNNPEWLQVDFQTKMTKVTGTTQGVKSLLSMYVKEFLISSQDGHQWTLFFQNGKVKV 1333
 QY 61 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 91
 DB 1334 FQGNQDSFTPVVNSLDPLLRILRIHPQSW 1364
 RESULT 12
 AAW33229
 ID AAW33229 standard; protein; 1383 AA.
 XX AAW33229;
 XX 21-OCT-2004 (revised)
 DT 30-APR-1998 (first entry)
 XX Procoagulant-active human factor VIII:C (FVIII) mutant protein.
 DE Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Region 1. .346
 FT /note= "factor VIIIA heavy chain"
 CC

FT Domain 1. .329
 FT /note= "A1 domain"
 FT Domain 1. .179
 FT /note= "plastocyanin-like domain 1"
 FT Disulfide-bond 153. .179
 FT /note= "probable"
 FT Domain 187. .329
 FT /note= "plastocyanin-like domain 2"
 FT Misc-difference 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT Cleavage-site 372. .373
 FT /note= "by thrombin"
 FT Domain 380. .711
 FT /note= "A2 domain"
 FT Domain 380. .554
 FT /note= "plastocyanin-like domain 3"
 FT Disulfide-bond 528. .554
 FT /note= "probable"
 FT Domain 564. .711
 FT /note= "plastocyanin-like domain 4"
 FT Region 711. .746
 FT /note= "a spacer of the sequence
 FT SFQNSRPSRTRQKFNATTPENDIEKTPWF AHRTPMPKIQNVSSDLMMLL
 FT is inserted between domains A2 and A3"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 FT Region 741. .1383
 FT /note= "factor VIIIA light chain"
 FT Domain 746. .1073
 FT /note= "A3 domain"
 FT Domain 1073. .1221
 FT /note= "C1 domain"
 FT Domain 1226. .1378
 FT /note= "C2 domain"
 FT XX
 XX WO3740145-A1.
 XX 30-OCT-1997.
 XX 24-APR-1997; 97WO-US0006563.
 XX 24-APR-1996; 96US-0016117P.
 XX 15-MAY-1996; 96US-0017785P.
 XX (UNWI) UNIV MICHIGAN.
 XX Kaufman RJ, Pipe SW, Amano K;
 XX WPI; 1997-535830/49.
 XX Modified human pro-coagulant active factor VIII - can be administered to
 XX haemophiliacs, i.e. factor VIII replacement therapy.
 XX Claim 19; Page; 57pp; English.
 XX The present sequence represents a novel pro-coagulant active factor VIII
 CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 CC Willebrand factor binding site, mutations F309S, R740A and addition of an
 CC amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
 CC along with calcium and phospholipid, acts as a cofactor for factor IXa,
 CC when it converts factor X to the activated form (factor Xa). FVIII is the
 CC coagulation factor deficient in the X-chromosome-linked bleeding disorder
 CC haemophilia A. Several other mutant FVIII proteins have also been created
 CC (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
 CC recombinant secretion at higher levels than typically obtained with wild
 CC type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
 CC (AAW33222) and R562K (AAW33223) are resistant to activated protein C
 CC (APC) cleavage. The present FVIII mutant can form a more stable
 CC configuration, and have an approximate 5-fold increase in specific
 CC activity compared to purified wild type FVIII, while increasing their
 CC binding affinity to von Willebrand factor improves their stability. The

CC FVIII proteins can be administered to haemophilias, i.e. FVIII
 CC replacement therapy, while the nucleic acid molecule can be used for gene
 CC therapy. note: this sequence does not appear in the specification; it was
 CC created using sequences from the given references
 CC
 CC Revised record issued on 21-OCT-2004 : Correction to feature table key
 CC
 CC Sequence 1383 AA;
 CC
 CC Query Match 100.0%; Score 484; DB 2; Length 1383;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-50;
 CC Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
 CC Db 1274 VNNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1333
 CC
 CC QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIRHPQSW 91
 CC Db 1334 FQGNQDSFTPVVNSLDPPLTRYLRIRHPQSW 1364
 CC
 CC RESULT 13
 CC ID ADQ37613
 CC ADQ37613 standard; protein; 1421 AA.
 CC AC ADQ37613;
 CC DT 21-OCT-2004 (first entry)
 CC DE Human Factor VIII deletion Pro739Phe mutant dB747.
 CC KW Human; Factor VIII; clotting factor; blood; blood disorder;
 CC KW haemophilia A; gene therapy; mutant; mutein; B domain.
 CC XX Homo sapiens.
 CC OS Synthetic.
 CC
 CC Key Location/Qualifiers
 CC Misc-difference 739 /note= "Wild-type Pro substituted by Phe"
 CC Misc-difference 747.748 /note= "Amino acids 748-1658 of the wild-type sequence
 CC have been deleted"
 CC
 CC US2004147436-A1.
 CC 29-JUL-2004.
 CC 28-JAN-2003; 2003US-00353753.
 CC 28-JAN-2003; 2003US-00353753.
 CC (KIMH/) KIM H.
 CC (SONG/) SONG I.
 CC (CHOI/) CHOI J W.
 CC (JANG/) JANG J.
 CC (KIMY/) KIM Y.
 CC (LEE/) LEE H S.
 CC (BANG/) BANG Y.
 CC (KIMD/) KIM D.
 CC Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
 CC WPT; 2004-561406/54.
 CC New Factor VIII polypeptides that are more stable than full-length Factor
 CC VIII, useful for clotting blood or for preventing or treating blood
 CC disorders, such as Hemophilia A.
 CC Example 4; Page; 39pp; English.
 CC The invention relates to a Factor VIII polypeptide comprising an internal

CC deletion of one or more amino acids between 1649 and 1688 fused to any
 CC amino acid sequence in B domain from about 741-782, with reference to
 CC full-length human Factor VIII appearing as ADQ37561. Also included are a
 CC pharmaceutical carrier, a lyophilised composition comprising the Factor
 CC VIII polypeptide, clotting blood in a subject (comprising contacting a
 CC clotting amount of the Factor VIII polypeptide with the blood), treating
 CC Haemophilia A in a patient (comprising administering a clotting amount of
 CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
 CC encoding the Factor VIII polypeptide, an expression vector comprising the
 CC nucleic acid operably linked to a promoter, a host cell comprising the
 CC expression vector, making the Factor VIII polypeptide (comprising
 CC culturing the cell in conditions suitable for the vector to express the
 CC polypeptide, and isolating the polypeptide) and a purified antibody
 CC specific for the Factor VIII polypeptide cited above. The internal
 CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
 CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
 CC -1658, 776-1658 or 783-1658. The Factor VIII polypeptide is a single
 CC chain. The proline at 739 is replaced by another amino acid. The
 CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
 CC encompassing fusion sites between Asn amino acid at positions 745, 757 or
 CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
 CC reference to the full-length human Factor VIII amino acid sequence cited
 CC above. The polypeptide comprises a formula with the following linked
 CC domains: H-S-L, where the H domain represents a polypeptide sequence
 CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
 CC the S domain represents a polypeptide spacing linker comprising up to
 CC about 60 amino acids and the L domain represents a polypeptide sequence
 CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
 CC composition and methods are useful for clotting blood or for preventing
 CC or treating blood disorders, such as Haemophilia A. The present sequence
 CC is a human factor VIII B domain deletion mutant also carrying the
 CC Pro739Phe mutation. NOTE: The present sequence is not shown in the
 CC specification but was created by the indexer using information in the
 CC specification.
 CC XX Sequence 1421 AA;
 CC SQ
 CC Query Match 100.0%; Score 484; DB 8; Length 1421;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-50;
 CC Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
 CC Db 1312 VNNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1371
 CC
 CC QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIRHPQSW 91
 CC Db 1372 FQGNQDSFTPVVNSLDPPLTRYLRIRHPQSW 1402
 CC
 CC RESULT 14
 CC ADQ37598
 CC ID ADQ37598 standard; protein; 1421 AA.
 CC AC ADQ37598;
 CC DT 21-OCT-2004 (first entry)
 CC DE Human Factor VIII deletion mutant dB747.
 CC KW Human; Factor VIII; clotting factor; blood; blood disorder;
 CC KW haemophilia A; gene therapy; mutant; mutein; B domain.
 CC XX Homo sapiens.
 CC OS Synthetic.
 CC
 CC Key Location/Qualifiers
 CC Misc-difference 747.748 /note= "Amino acids 748-1658 of the wild-type sequence
 CC have been deleted"
 CC
 CC US2004147436-A1.

XX PD 29-JUL-2004.
XX PF 28-JAN-2003; 2003US-00353753.
XX PR 28-JAN-2003; 2003US-00353753.
XX PA (KIMH//) KIM H.
XX PA (SONG//) SONG I.
XX PA (CHOI//) CHOI J. W.
XX PA (JANG//) JANG J.
XX PA (KIMY//) KIM Y.
XX PA (LEE//) LEE H. S.
XX PA (BANG//) BANG Y.
XX PA (KIMD//) KIM D.
XX PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
XX DR WPI; 2004-561406/54.
XX
XX New Factor VIII polypeptides that are more stable than full-length Factor VIII, useful for clotting blood or for preventing or treating blood disorders, such as Hemophilia A.
XX PT
XX PS Claim 2; Page; 39pp; English.
XX
XX The invention relates to a Factor VIII polypeptide comprising an internal deletion of one or more amino acids between 1649 and 1688 fused to any amino acid sequence in B domain from about 741-782, with reference to full-length human Factor VIII appearing as ADQ37564. Also included are a pharmaceutical composition comprising the Factor VIII polypeptide (and a pharmaceutical carrier), a lyophilised composition comprising the Factor VIII polypeptide, clotting blood in a subject (comprising contacting a clotting amount of the Factor VIII polypeptide with the blood), treating Haemophilia A in a patient (comprising administering a clotting amount of the above Factor VIII polypeptide to a patient), an isolated nucleic acid encoding the Factor VIII polypeptide, an expression vector comprising the nucleic acid operably linked to a promoter, a host cell comprising the expression vector, making the Factor VIII polypeptide (comprising culturing the cell in conditions suitable for the vector to express the polypeptide, and isolating the polypeptide) and a purified antibody specific for the Factor VIII polypeptide cited above. The internal deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652, 758-1655, 765-1649, 765-1652, 765-1655, 768-1658, 769 chain. The proline at 739 is replaced by another amino acid. The polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser) encompassing fusion sites between Asn amino acid at positions 745, 757 or 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with reference to the full-length human Factor VIII amino acid sequence cited above. The polypeptide comprises a formula with the following linked domains: H-S-L, where the H domain represents a polypeptide sequence comprising substantially Ala-1 through Arg-740 of the human Factor VIII, the S domain represents a polypeptide spacing linker comprising up to about 60 amino acids and the L domain represents a polypeptide sequence comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The composition and methods are useful for clotting blood or for preventing or treating blood disorders, such as Haemophilia A. The present sequence is a human factor VIII B domain deletion mutant. NOTE: The present sequence is not shown in the specification but was created by the indexer using information in the specification.
XX SQ Sequence 1421 AA;
Query Match 100.0%; Score 484; DB 8; Length 1421;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNNPEKWLQVDFTKMTKVTGTTQGVKSLTSMYVKEFLISSQDGHOWTLFFQNGKVKV 60
DB 1312 VNNPEKWLQVDFTKMTKVTGTTQGVKSLTSMYVKEFLISSQDGHOWTLFFQNGKVKV 1371
QY 61 FQGNQDSFTPVVNSLDPLLRHLPQSW 91

DB 1372 FQGNQDSFTPVVNSLDPLLRHLPQSW 1402
RESULT 15
ADQ37591
XX ADQ37591 standard; protein; 1422 AA.
XX AC ADQ37591;
XX DT 21-OCT-2004 (first entry)
XX DE Human Factor VIII Glycosylation/deletion mutant dBN(45-56).
XX KW Human; Factor VIII; clotting factor; blood; blood disorder;
XX KW haemophilia A; gene therapy; mutant; mutein; N-glycosylation.
XX OS Homo sapiens.
XX OS Synthetic.
XX Key Location/Qualifiers
XX FT Modified-site 745..747 /note= "Asn is N-glycosylated"
XX FT Misc-difference 745..746 /note= "Amino acids 746-1655 of the wild-type sequence have been deleted"
XX FT
XX US2004147436-A1.
XX PN 29-JUL-2004.
XX PD 28-JAN-2003; 2003US-00353753.
XX PF 28-JAN-2003; 2003US-00353753.
XX PR 28-JAN-2003; 2003US-00353753.
XX PA (KIMH//) KIM H.
XX PA (SONG//) SONG I.
XX PA (CHOI//) CHOI J. W.
XX PA (JANG//) JANG J.
XX PA (KIMY//) KIM Y.
XX PA (LEE//) LEE H. S.
XX PA (BANG//) BANG Y.
XX PA (KIMD//) KIM D.
XX PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
XX DR WPI; 2004-561406/54.
XX New Factor VIII polypeptides that are more stable than full-length Factor VIII, useful for clotting blood or for preventing or treating blood disorders, such as Hemophilia A.
XX PS Claim 2; Page; 39pp; English.
XX The invention relates to a Factor VIII polypeptide comprising an internal deletion of one or more amino acids between 1649 and 1688 fused to any amino acid sequence in B domain from about 741-782, with reference to full-length human Factor VIII appearing as ADQ37564. Also included are a pharmaceutical composition comprising the Factor VIII polypeptide (and a pharmaceutical carrier), a lyophilised composition comprising the Factor VIII polypeptide, clotting blood in a subject (comprising contacting a clotting amount of the Factor VIII polypeptide with the blood), treating Haemophilia A in a patient (comprising administering a clotting amount of the above Factor VIII polypeptide to a patient), an isolated nucleic acid encoding the Factor VIII polypeptide, an expression vector comprising the nucleic acid operably linked to a promoter, a host cell comprising the expression vector, making the Factor VIII polypeptide (comprising culturing the cell in conditions suitable for the vector to express the polypeptide, and isolating the polypeptide) and a purified antibody specific for the Factor VIII polypeptide cited above. The internal deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652, 758-1655, 765-1649, 765-1652, 765-1655, 768-1658, 769 chain. The proline at 739 is replaced by another amino acid. The polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser) encompassing fusion sites between Asn amino acid at positions 745, 757 or 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with reference to the full-length human Factor VIII amino acid sequence cited above. The polypeptide comprises a formula with the following linked domains: H-S-L, where the H domain represents a polypeptide sequence comprising substantially Ala-1 through Arg-740 of the human Factor VIII, the S domain represents a polypeptide spacing linker comprising up to about 60 amino acids and the L domain represents a polypeptide sequence comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The composition and methods are useful for clotting blood or for preventing or treating blood disorders, such as Haemophilia A. The present sequence is a human factor VIII B domain deletion mutant. NOTE: The present sequence is not shown in the specification but was created by the indexer using information in the specification.

CC chain. The proline at 739 is replaced by another amino acid. The
CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC comprising fusion sites between Asn amino acid at positions 745,757 or
CC 767, and the Ser amino acid at positions 1651, 1654 or 1657, with
CC reference to the full length human factor VIII amino acid sequence cited
CC above. The polypeptide comprises a formula with the following linked
CC domains: H-81, peptide comprising the H domain, represents a polypeptide sequence
CC comprising substantially Ala1 through Arg-740 of the human factor VIII,
CC the S domain, represents a polypeptide spacing linker comprising up to
CC about 60 amino acids and the L domain, represents a polypeptide sequence
CC comprising Arg-1649 through Tyr-3231 of the human factor VIII. The
CC co-siting and method are useful for clotting blood or for preventing
CC or treating blood disorders such as Hemophilia A. The present sequence
CC is a human factor VIII deletion mutant where the deletion creates an N-
CC glycosylation site. NOTE: The present sequence is not shown in the
CC specification but was created by the indexer using information in the
CC specification.

XX
SQ Sequence 1422 AA;

Query Match 100.0%; Score 484; DB 8; Length 1422;
Best Local Similarity 100.0%; Pred No. 1.8e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPKEWLQVDFQKTMKVTGTTGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKV 60
Db 1313 VNNPKEWLQVDFQKTMKVTGTTGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKV 1372
Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRHPQSW 91
Db 1373 FQGNQDSFTPVVNSLDPPLLTRYLRHPQSW 1403

Search completed: April 20, 2005, 17:19:11
Job time : 28.3812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:14:28 ; Search time 8.78465 Seconds
(without alignments)
773.288 Million cell updates/sec

Title: US-09-853-080b-49_COPY_2242_2332

Perfect score: 484

Sequence: 1 VNNPKWLQVDFQTKMKTG.....VNSLDPPLLRYLRIHQSW 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	100.0	109	1	US-08-111-939-21
2	484	100.0	160	2	US-08-162-402B-15
3	484	100.0	160	4	US-09-331-793-9
4	484	100.0	208	4	US-09-949-016-11661
5	484	100.0	1438	3	US-09-209-916-1
6	484	100.0	1457	4	US-09-001-039B-47
7	484	100.0	1471	1	US-08-683-839B-3
8	484	100.0	1661	2	US-08-882-083-2
9	484	100.0	1661	2	US-08-558-107-2
10	484	100.0	1661	3	US-09-243-539-2
11	484	100.0	2332	1	US-07-864-004B-4
12	484	100.0	2332	1	US-08-251-937A-4
13	484	100.0	2332	1	US-08-212-133A-2
14	484	100.0	2332	1	US-08-276-594A-2
15	484	100.0	2332	1	US-08-474-503-2
16	484	100.0	2332	2	US-08-670-707A-2
17	484	100.0	2332	3	US-09-037-601-2
18	484	100.0	2332	3	US-09-324-867-3
19	484	100.0	2332	3	US-09-315-179-2
20	484	100.0	2332	4	US-09-523-656-2
21	484	100.0	2332	4	US-09-957-641A-2
22	484	100.0	2332	5	PCT-US93-03275-4
23	484	100.0	2332	5	PCT-US94-13200-2
24	484	100.0	2351	1	US-08-121-202-2
25	484	100.0	2351	1	US-08-366-851A-2
26	484	100.0	2351	4	US-10-133-907-4
27	484	100.0	2351	4	US-09-001-039B-45

28	484	100.0	2351	6	5171844-2	Patent No. 5171844
29	484	100.0	2351	6	5171844-2	Patent No. 5171844
30	480	99.2	218	1	US-07-607-538C-5	Sequence 5, Appli
31	480	99.2	218	2	US-08-162-402B-5	Sequence 5, Appli
32	480	99.2	218	4	US-09-364-185-5	Sequence 5, Appli
33	475	98.1	2351	6	5422260-1	Patent No. 5422260
34	475	98.1	2351	6	5422260-1	Patent No. 5422260
35	419	86.6	130	4	US-09-331-793-1	Sequence 1, Appli
36	417	86.2	2343	3	US-09-324-867-2	Sequence 2, Appli
37	406	83.9	160	4	US-09-957-641A-20	Sequence 20, Appli
38	406	83.9	2304	3	US-09-324-867-4	Sequence 4, Appli
39	406	83.9	2319	1	US-08-212-133A-8	Sequence 8, Appli
40	406	83.9	2319	1	US-08-474-503-6	Sequence 6, Appli
41	406	83.9	2319	2	US-08-670-707A-6	Sequence 6, Appli
42	406	83.9	2319	3	US-09-037-601-6	Sequence 6, Appli
43	406	83.9	2319	3	US-09-315-179-6	Sequence 6, Appli
44	406	83.9	2319	4	US-09-523-656-28	Sequence 28, Appli
45	406	83.9	2319	5	PCT-US94-13200-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-111-939-21
; Sequence 21, Application US/08111939
; Patent No. 5460951
; GENERAL INFORMATION:
; APPLICANT: Kawai, Shinji
; APPLICANT: Takeehita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Amanu, Egon
; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like
; TITLE OF INVENTION: Protein and Process for its Production
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/111.939
; FILING DATE: 26-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 324033/92
; FILING DATE: 03-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 230029/92
; FILING DATE: 28-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Portman, David S.
; REGISTRATION NUMBER: 33,694
; REFERENCE/DOCKET NUMBER: 02481.1321-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4000
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-111-939-21

Query Match 100.0%; Score 484; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 5,7e-54;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 6 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 65

Qy 61 FQGNQDSFTPVVNSLDPLLRIRHPQSW 91
Db 66 FQGNQDSFTPVVNSLDPLLRIRHPQSW 96

RESULT 2
US-08-162-402B-15
; Sequence 15, Application US/08162402B
; Patent No. 5972337
; GENERAL INFORMATION:
; APPLICANT: CERIANI, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; APPLICANT: LAROCCA, DAVID J.
; TITLE OF INVENTION: 46 KDALTON HUMAN MILK PAT
; TITLE OF INVENTION: GLOBULE (HMGF) ANTIGEN, FRAGMENTS & FUSION PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder & Poplawski
; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Anzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-15

Query Match 100.0%; Score 484; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 9,6e-54;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 51 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 110

Qy 61 FQGNQDSFTPVVNSLDPLLRIRHPQSW 91
Db 111 FQGNQDSFTPVVNSLDPLLRIRHPQSW 141

RESULT 3
US-09-331-793-9
; Sequence 9, Application US/09331793
; Patent No. 6500646
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi
; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/09/331,793
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-331-793-9

Query Match 100.0%; Score 484; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 9,6e-54;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 51 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 110

Qy 61 FQGNQDSFTPVVNSLDPLLRIRHPQSW 91
Db 111 FQGNQDSFTPVVNSLDPLLRIRHPQSW 141

RESULT 4
US-09-949-016-11661
; Sequence 11661, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11661
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11661

Query Match 100.0%; Score 484; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 1,4e-53;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 99 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 158

Qy 61 FQGNQDSFTPVVNSLDPLLRIRHPQSW 91
Db 159 FQGNQDSFTPVVNSLDPLLRIRHPQSW 189

RESULT 5
US-09-209-916-1
; Sequence 1, Application US/09209916

Patent No. 6358703
GENERAL INFORMATION:
APPLICANT: Cho, Myung-Sam
APPLICANT: Chan, Sham-Yuen
APPLICANT: Kelsey, William
APPLICANT: Yee, Helena
TITLE OF INVENTION: Expression System for Factor VIII
FILE REFERENCE: MSB-7255
CURRENT APPLICATION NUMBER: US/09/209,916
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1438
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from human factor VIII sequence
US-09-209-916-1

Query Match 100.0%; Score 484; DB 3; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.9e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKWLQVDFQKTMKVTVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1388

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 1419

RESULT 6
US-09-001-039B-47
Sequence 47, Application US/09001039B
Patent No. 6818439
GENERAL INFORMATION:
APPLICANT: Jolly, Douglas J.
APPLICANT: Chang, Stephen M.W.
APPLICANT: Respass, James G.
APPLICANT: DePolo, Nicholas J.
APPLICANT: Hsu, David Chi-Tang
APPLICANT: Ibanez, Carlos E.
APPLICANT: Greengard, Judith
APPLICANT: Lee, Will
TITLE OF INVENTION: METHODS FOR ADMINISTRATION OF
TITLE OF INVENTION: RECOMBINANTGENE DELIVERY VEHICLES FOR TREATMENT
TITLE OF INVENTION: OF HEMOPHILIA AND OTHER DISORDERS
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001,039B
FILING DATE: 13-JAN-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 1155.005 / 930049.441C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

Patent No. 682-6031
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 1457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-001-039B-47

Query Match 100.0%; Score 484; DB 4; Length 1457;
Best Local Similarity 100.0%; Pred. No. 2e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1348 VNNPKWLQVDFQKTMKVTVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1407

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 1408 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 1438

RESULT 7
US-08-683-839B-3
Sequence 3, Application US/08683839B
Patent No. 5744326
GENERAL INFORMATION:
APPLICANT: Ili, Charles R. et al.
TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
TITLE OF INVENTION: Regulatory Sequences to Increase Expression of
TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,839B
FILING DATE: 11-MARCH-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Remillard, Jane E.
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: TTI-138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-839B-3

Query Match 100.0%; Score 484; DB 1; Length 1471;
Best Local Similarity 100.0%; Pred. No. 2e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKWLQVDFQKTMKVTVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60

Db 1362 VNNPKWLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1421
QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1422 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1452

RESULT 8

US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-882-083-2

Query Match 100.0%; Score 484; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.4e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNNPKWLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1552 VNNPKWLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1611
QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1612 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1642

RESULT 9

US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-107-2

Query Match 100.0%; Score 484; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.4e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNNPKWLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1552 VNNPKWLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1611
QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1612 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1642

RESULT 10

US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715


```

; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-243-539-2

Query Match 100.0%; Score 484; DB 3; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.4e-52; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1552 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1611

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 1612 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 1642

RESULT 11
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/864,004B
; FILING DATE: 07 APRIL 1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
US-07-864-004B-4

Query Match 100.0%; Score 484; DB 3; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.4e-52; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1552 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1611

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 1612 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 1642

RESULT 12
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
US-08-251-937A-4

Query Match 100.0%; Score 484; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 3.7e-52; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 2223 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 2282

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 2283 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 2313

RESULT 12
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
US-08-251-937A-4

Query Match 100.0%; Score 484; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 3.7e-52; Indels 0; Gaps 0;
Matches 91; Conservative 0; Mismatches 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 2223 VNNPKEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 2282

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 91
Db 2283 FQGNQDSFTPVVNSLDPPLTRYLRHPQSW 2313
```

```

Db      2283 FQGNQDSFTPVVNSLDPLLRHLPQSW 2313

RESULT 13
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Ruge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabsst, Patricia L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-08-212-133A-2
Query Match 100.0%; Score 484; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 3.7e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db      2223 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 2282

Qy      61 FQGNQDSFTPVVNSLDPLLRHLPQSW 91
Db      2283 FQGNQDSFTPVVNSLDPLLRHLPQSW 2313

RESULT 14
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka

```

```

; APPLICANT: SUGAWARA, Keishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-594A-2
Query Match 100.0%; Score 484; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 3.7e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db      2223 VNNPKWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 2282

Qy      61 FQGNQDSFTPVVNSLDPLLRHLPQSW 91
Db      2283 FQGNQDSFTPVVNSLDPLLRHLPQSW 2313

RESULT 15
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU06CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
;
US-08-474-503-2

Query Match 100.0%; Score 484; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 3.7e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTGQVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKV 60
Db 2223 VNNPKEWLQVDFQKTMKVTGVTGQVKSLTSMYVKEFLISSODGHQWTLFFQNGKVKV 2282

Qy 61 FQGNQDSFTPVNSLDPPLLTRYLRHPSW 91
Db 2283 FQGNQDSFTPVNSLDPPLLTRYLRHPSW 2313
```

Search completed: April 20, 2005, 17:22:32
Job time : 8.78465 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:21:09 ; Search time 19.5965 Seconds
(without alignments)
1543.422 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332

Perfect score: 484

Sequence: 1 VNNPKEWLQVDFQTKMTKVTG.....VNSLDPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	100.0	160	15	US-10-298-796-9
2	484	100.0	216	17	US-10-741-600-1033
3	484	100.0	1438	13	US-10-006-091-1
4	484	100.0	1438	13	US-10-047-257-1
5	484	100.0	1438	14	US-10-225-900-1
6	484	100.0	1459	15	US-10-239-498A-13
7	484	100.0	1471	13	US-10-095-718-2
8	484	100.0	1471	15	US-10-681-970-2
9	484	100.0	2332	9	US-09-957-641-2
10	484	100.0	2332	14	US-10-187-319-2
11	484	100.0	2332	14	US-10-131-510A-2
12	484	100.0	2332	15	US-10-445-235-2
13	484	100.0	2332	15	US-10-360-101-229

14	484	100.0	2332	15	US-10-239-498A-2	Sequence 2, Appli
15	484	100.0	2332	16	US-10-466-998A-1	Sequence 1, Appli
16	484	100.0	2332	16	US-10-721-997A-34	Sequence 34, Appli
17	484	100.0	2351	14	US-10-132-829-4	Sequence 4, Appli
18	484	100.0	2351	14	US-10-172-712-27	Sequence 27, Appli
19	484	100.0	2351	14	US-10-133-907-4	Sequence 4, Appli
20	484	100.0	2351	15	US-10-411-037-30	Sequence 30, Appli
21	484	100.0	2351	15	US-10-411-026-30	Sequence 30, Appli
22	484	100.0	2351	15	US-10-410-962-30	Sequence 30, Appli
23	484	100.0	2351	15	US-10-411-049-30	Sequence 30, Appli
24	484	100.0	2351	16	US-10-410-930-30	Sequence 30, Appli
25	484	100.0	2351	16	US-10-410-997-30	Sequence 30, Appli
26	484	100.0	2351	16	US-10-411-012-30	Sequence 30, Appli
27	484	100.0	2351	16	US-10-287-994-30	Sequence 30, Appli
28	484	100.0	2351	16	US-10-410-913-30	Sequence 30, Appli
29	484	100.0	2351	17	US-10-741-600-1034	Sequence 1034, Ap
30	484	100.0	2351	17	US-10-410-980-30	Sequence 30, Appli
31	480	99.2	218	15	US-10-038-252-5	Sequence 5, Appli
32	480	99.2	1459	15	US-10-239-498A-4	Sequence 4, Appli
33	480	99.2	1459	15	US-10-239-498A-15	Sequence 15, Appli
34	419	86.6	130	15	US-10-298-796-1	Sequence 1, Appli
35	417	86.2	1431	13	US-10-095-718-4	Sequence 4, Appli
36	417	86.2	1431	15	US-10-681-970-4	Sequence 4, Appli
37	406	83.9	2319	14	US-10-187-319-6	GENERAL INFORMA
38	406	83.9	2319	14	US-10-131-510A-6	Sequence 6, Appli
39	392	81.0	1443	14	US-10-187-319-39	Sequence 39, Appli
40	392	81.0	1443	14	US-10-131-510A-39	Sequence 39, Appli
41	392	81.0	2114	16	US-10-721-997A-33	Sequence 33, Appli
42	392	81.0	2133	14	US-10-187-319-37	Sequence 37, Appli
43	392	81.0	2133	14	US-10-131-510A-37	Sequence 37, Appli
44	222	45.9	60	15	US-10-298-796-8	Sequence 8, Appli
45	207	42.8	434	16	US-10-485-360-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-10-298-796-9
; Sequence 9, Application US/10298796
; Publication No. US20030220490A1
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi
; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/10/298,796
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US/09/331,793
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-298-796-9

Query Match 100.0%; Score 484; DB 15; Length 160;

Best Local Similarity 100.0%; Pred. No. 9.9e-50; Indels 0; Gaps 0;

Matches 91; Conservative 0; Mismatches 0;

Qy	1	VNNPKEWLQVDFQTKMTKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV	60
Db	51	VNNPKEWLQVDFQTKMTKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV	110

Qy 61 FQGNQDSFTPVNSLDPLLTRYLRIHPQSW 91

Db 111 FQGNQDSFTPVNSLDPLLTRYLRIHPQSW 141

RESULT 2

US-10-741-600-1033

```

; Sequence 1033, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 1033
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-741-600-1033

Query Match 100.0%; Score 484; DB 17; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.4e-49;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 107 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 166

QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 167 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 197

RESULT 3
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match 100.0%; Score 484; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1388

QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1419

RESULT 4
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam

```

```

; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match 100.0%; Score 484; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1388

QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1419

RESULT 5
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US2003007752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

Query Match 100.0%; Score 484; DB 14; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1388

QY 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1419

RESULT 6
US-10-239-498A-13
; Sequence 13, Application US/10239498A

```

Publication No. US2004002333A1
GENERAL INFORMATION:
APPLICANT: Hauser, Charlotte
APPLICANT: Horster, Andrea
APPLICANT: Schroder, Carola
APPLICANT: Lehner, Michael
TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in Human Cell Lines
FILE REFERENCE: 80977.0001
CURRENT APPLICATION NUMBER: US/10/239,498A
PRIOR FILING DATE: 2003-07-08
PRIOR APPLICATION NUMBER: PCT/EP01/03220
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 1459
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PTGF8-2hyg-s
US-10-239-498A-13
Query Match 100.0%; Score 484; DB 15; Length 1459;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1350 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1409
Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1410 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1440
RESULT 7
US-10-095-718-2
Sequence 2, Application US/10095718
Publication No. US20020131956A1
GENERAL INFORMATION:
APPLICANT: Walsh, Christopher
APPLICANT: Chao, Hengjun
APPLICANT: Burstein, Haim
APPLICANT: Lynch, Carmel
APPLICANT: Stepan, Tony
APPLICANT: Munson, Keith
TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same
FILE REFERENCE: 35052/204375
CURRENT APPLICATION NUMBER: US/10/095,718
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 09/689,430
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/158,780
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1471
TYPE: PRT
ORGANISM: Homo sapiens B-domain deleted factor VIII
FEATURE:
OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2
Query Match 100.0%; Score 484; DB 13; Length 1471;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1362 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1421

Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1422 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1452
RESULT 8
US-10-681-970-2
Sequence 2, Application US/10681970
Publication No. US20040062752A1
GENERAL INFORMATION:
APPLICANT: Walsh, Christopher
APPLICANT: Chao, Hengjun
APPLICANT: Burstein, Haim
APPLICANT: Lynch, Carmel
APPLICANT: Stepan, Tony
APPLICANT: Munson, Keith
TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same
FILE REFERENCE: 35052/204375
CURRENT APPLICATION NUMBER: US/10/681,970
PRIOR FILING DATE: 2003-10-09
PRIOR APPLICATION NUMBER: US/09/689,430
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/158,780
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1471
TYPE: PRT
ORGANISM: Homo sapiens B-domain deleted factor VIII
FEATURE:
OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-681-970-2
Query Match 100.0%; Score 484; DB 15; Length 1471;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 1362 VNNPEWLQVDFQKTMKVTGTTQGVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 1421
Qy 61 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 91
Db 1422 FQGNQDSFTPVVNSLDPPLTRYLRIHPQSW 1452
RESULT 9
US-09-957-641-2
Sequence 2, Application US/09957641
Publication No. US20020182670A1
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: MODIFIED FACTOR VIII
FILE REFERENCE: 75-00
CURRENT APPLICATION NUMBER: US/09/957,641
CURRENT FILING DATE: 2001-09-16
PRIOR APPLICATION NUMBER: US 60/234047
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: US 60/236460
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-641-2
Query Match 100.0%; Score 484; DB 9; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.8e-48;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
 Db 2223 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 2282
 QY 61 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 91
 Db 2283 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 2313

RESULT 10

US-10-187-319-2
 ; Sequence 2, Application US/10197319
 ; Publication No. US20030068785A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/187,319
 ; FILING DATE: 27-Aug-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/523,656
 ; FILING DATE: 2000-03-10
 ; APPLICATION NUMBER: US 09/037,601
 ; FILING DATE: 1998-03-10
 ; APPLICATION NUMBER: WO PCT/US97/11155
 ; FILING DATE: 1997-06-26
 ; APPLICATION NUMBER: US 08/670,707
 ; FILING DATE: 1996-06-26
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Greenlee, Lorraine L.
 ; REGISTRATION NUMBER: 27,894
 ; REFERENCE/DOCKET NUMBER: 75-95K
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/499-8080
 ; TELEFAX: 303/499-8089
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2332 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; TISSUE TYPE: Liver
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-187-319-2

Query Match 100.0%; Score 484; DB 14; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 2.8e-48;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
 Db 2223 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 2282

Db 2223 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 2282

QY 61 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 91
 Db 2283 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 2313

RESULT 11

US-10-131-510A-2
 ; Sequence 2, Application US/10131510A
 ; Publication No. US20030166536A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Modified Factor VIII
 ; FILE REFERENCE: 75-95J
 ; CURRENT APPLICATION NUMBER: US/10/131,510A
 ; CURRENT FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: U.S. 09/315,179
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: U.S. 09/037,601
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: U.S. 08/670,707
 ; PRIOR FILING DATE: 1996-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US97/11155
 ; PRIOR FILING DATE: 1997-06-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/13200
 ; PRIOR FILING DATE: 1994-11-15
 ; PRIOR APPLICATION NUMBER: U.S. 08/212,133
 ; PRIOR FILING DATE: 1994-03-11
 ; PRIOR APPLICATION NUMBER: U.S. 07/864,004
 ; PRIOR FILING DATE: 1992-04-07
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 2332
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens

Query Match 100.0%; Score 484; DB 14; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 2.8e-48;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
 Db 2223 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 2282

QY 61 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 91
 Db 2283 FQGNQDSFTPVVNSLDPPLLTTRYLRHPQSW 2313

RESULT 12

US-10-445-235-2
 ; Sequence 2, Application US/10445235
 ; Publication No. US20040005670A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Katherine A. High
 ; APPLICANT: Rodney M. Camire
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; TREATMENT OF HEMOPHILIA A
 ; FILE REFERENCE: CHOP0176
 ; CURRENT APPLICATION NUMBER: US/10/445,235
 ; CURRENT FILING DATE: 2003-05-22
 ; PRIOR APPLICATION NUMBER: 60/382,486
 ; PRIOR FILING DATE: 2002-05-22
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 2332
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; ORGANISM: homo sapiens

Query Match 100.0%; Score 484; DB 14; Length 2332;
 Best Local Similarity 100.0%; Pred. No. 2.8e-48;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 60
 Db 2223 VNNPKWLQVDFQKTMKVGTGGVKSLLTSMYKFLISSQDGHQWTLFFQNGKVKV 2282

Query March	100.0%	Score 484;	DB 15;	Length 2332;
Best Local Similarity	100.0%;	Prod. No. 2.8e-48;		
Matches 91;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy.	1	VNNPEWLQVDFQTKMKVTGTTGCVKSLTSMYVKEFLTISSSDGHQWTLFFQNGVKV	60	
Db	2223	VNNPEWLQVDFQTKMKVTGTTGCVKSLTSMYVKEFLTISSSDGHQWTLFFQNGVKV	2282	
Qy	61	FQGNQDSFTPVNSLDPPLLTRYLRHHPQS	91	
Db	2283	FOGNQDSFTPVNSLDPPLLTRYLRHHPQS	2313	

```

RESULT 13
US-10-360-101-229
; Sequence 229, Application US/10360101
; Publication No. US2004009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leshouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of factor VIII
US-10-360-101-229

```

Query Match	100.0%	Score 484;	DB 15;	Length 2332;
Best Local Similarity	100.0%;	Pred. No. 2.8e-48;		
Matches 91;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VNNPKWLQVD	FQKTMKVGTGTTQGVKSLTSMYVKEFLISSSDG	HQWTLFFQNGKVKV 60
Db	2223	VNNPKWLQVD	FQKTMKVGTGTTQGVKSLTSMYVKEFLISSSDG	HQWTLFFQNGKVKV 2382
Qy	61	FOGNQSF	TPVNSLDPPLLTRYLRIHPQSW	91
Db	2283	FOGNQSF	TPVNSLDPPLLTRYLRIHPQSW	2313

```

RESULT 14
US-10-239-498A-2
; Sequence 2, Application US/10239498A
; Publication No. US20040023333A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Charlotte
; APPLICANT: Horster, Andrea
; APPLICANT: Schroder, Carola
; APPLICANT: Lehnerer, Michael
; TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
; FILE REFERENCE: 80977.0001
; CURRENT APPLICATION NUMBER: US/10/239,498A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/EP01/03220
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-498A-2

```

US-10-239-498A-2

	Query Match	100.0%;	Score 484;	DB 15;	Length 2332;
	Best Local Similarity	100.0%;	Pred. No. 2.8e-48;		
	Matches 91;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VNNPKEWLQVDFQRTMKVTGVTGGVKSLTSMYVKFLLISSQDGHQWTLFFQNGKVKV	60		
Db	2223	VNNPKEWLQVDFQRTMKVTGVTGGVKSLTSMYVKFLLISSQDGHQWTLFFQNGKVKV	2282		
Qy	61	FOGQNSFTPVNNSLDPPLLTRVLRTHPQSW	91		
Db	2283	FOGQNSFTPVNNSLDPPLLTRVLRTHPQSW	2313		

RESULT 15
US-10-466-998A-1
; Sequence 1, Application US/10466998A
; Publication No. US20040126856A1
; GENERAL INFORMATION:
; APPLICANT: BAJAJ, S. Paul
; APPLICANT: FAY, Philip J.
; TITLE OF INVENTION: Factor IXa: Factor VIIIA Interaction and Methods
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 66153-41436
; CURRENT APPLICATION NUMBER: US/10/466,998A
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/01724
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 60/263,431
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-998A-1

Query Match	100.0%;	Score 484;	DB 16;	Length 2332;
Best Local Similarity	100.0%;	Pred. No. 2.8e-48;		
Matches 91;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VNNPKWLQVDFQTKMKVTGVTGGYKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV	60	
Db	2223	VNNPKWLQVDFQTKMKVTGVTGGYKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV	2282	
Qy	61	FOGNQDSFTEPVNNSLDPPLLTRYLRTHPSW	91	
Db	2283	FOGNQDSFTEPVNNSLDPPLLTRYLRTHPSW	2313	

Search completed: April 20, 2005, 17:34:49
Job time : 19.5965 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:13:08 ; Search time 6.30693 Seconds
(without alignments)
1388.270 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332

Perfect score: 484
Sequence: 1 VNNPKWLQVDFQTKMKVTG.....VNSLDPPLLTRYLRHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484	100.0	216	2 A44258	factor VIII-associ
2	484	100.0	2351	1 EZHU	coagulation factor
3	406	83.9	2319	2 A47004	coagulation factor
4	392	81.0	2133	2 T42763	coagulation factor
5	208	43.0	427	2 JC4915	age protein precu
6	207	42.8	463	1 A36479	milk fat globule m
7	206	42.6	2183	2 T42764	coagulation factor
8	199	41.1	2224	1 KFHU5	coagulation factor
9	198	40.9	2211	1 KFB05	coagulation factor
10	174	36.0	409	2 T11743	p247 protein - pig
11	174	36.0	845	2 JCS256	adipocyte transcr
12	173	35.7	719	2 S51739	transcription repr
13	164	33.9	218	2 A47285	milk fat globule p
14	159	32.9	401	2 S65138	Glycoprotein antig
15	159	32.9	427	2 S74211	PA5-6/7 protein pr
16	154	31.8	927	1 JQ0948	A5 antigen precurs
17	137.5	28.4	3133	2 S52093	hemocytin - silkw
18	127.5	26.3	737	2 T31349	hypothetical prote
19	123	25.4	1381	2 T31083	paranodin - rat
20	123	25.4	1385	2 T41558	neurexin IV - mous
21	115.5	23.9	737	2 T15615	hypothetical prote
22	112.5	23.2	913	2 A48280	receptor tyrosine k
23	111.5	23.0	876	2 A49508	protein-tyrosine k
24	107.5	22.2	910	2 A53137	tyrosine kinase re
25	104.5	21.6	1283	2 T13799	neurexin IV - frui
26	103.5	21.4	819	2 T48859	tyro 10 receptor k
27	98.5	20.4	855	2 S42821	protein-tyrosine k
28	83	17.1	791	2 T16031	hypothetical prote
29	78	16.1	491	2 T40455	penicillin binding

30	73	15.1	1130	2 A48843	MHC class II trans
31	72.5	15.0	392	2 JC7633	aryl hydrocarbon n
32	71.5	14.8	157	2 H69849	hypothetical prote
33	71.5	14.8	1082	2 T42204	chromatin structur
34	70.5	14.6	589	2 T35784	probable secreted
35	69.5	14.4	441	2 C38575	gluconate 2-dehydr
36	68.5	14.2	762	2 H71916	DNA mismatch repai
37	68.5	14.2	805	2 JC7635	aryl hydrocarbon r
38	68	14.0	334	2 H75362	hypothetical prote
39	68	14.0	512	2 T47554	cytochrome P450 ho
40	67.5	13.9	221	2 C64040	hypothetical prote
41	67.5	13.9	1134	2 T20332	hypothetical prote
42	67	13.8	404	2 JN0630	microbial aspartic
43	67	13.8	853	2 AC2079	ferrichrome-iron r
44	67	13.8	2825	2 T14271	Doc4 protein, stre
45	66.5	13.7	358	2 S58376	aryl hydrocarbon r

ALIGNMENTS

RESULT 1

A44258
factor VIII-associated gene B hypothetical protein - human
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A44258
R;Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-589, 1992
A:Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; MUID:93052386; PMID:1427887
A/Accession: A44258
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-216 <LV>
A/Cross-references: UNIPROT:Q14286; GB:M90707; NID:g182316; PIDN:AAA58466.1; PID:g182317
C/Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
F;57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 100.0%; Score 484; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.6e-45;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VNNPKWLQVDFQTKMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db	107	VNNPKWLQVDFQTKMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 166
Qy	61	FGNQDSFTPVNSLDPPLLTRYLRHPQSW 91
Db	167	FGNQDSFTPVNSLDPPLLTRYLRHPQSW 197

RESULT 2

EZHU
coagulation factor VIII precursor [validated] - human
N;Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant compor
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
R;Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.
A:Reference number: I54318; MUID:93265012; PMID:1303178
A/Accession: I54318
A/Status: preliminary; translated from GB/EMBL/DBDJB
A/Molecule type: DNA
A/Residues: 1-1921,'S',1923-2351 <RES>
A/Cross-references: UNIPROT:P00451; GB:M88648; NID:g182381; PIDN:AAA52420.1; PID:g182383
R;Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg, Nature 312, 330-337, 1984
A:Title: Expression of active human factor VIII from recombinant DNA clones.
A:Reference number: A00525; MUID:85061548; PMID:6438526

A:Accession: A00525
A:Molecule type: mRNA
A:Residues: 1-2351 <WO>
A:Cross-references: EMBL:X01166; EMBL:X01179
R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D.;
s, D.N.; Hewick, R.M.
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A:Reference number: 158059, PMID:85061550; PMID:6438528
A:Accession: I58059
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-74, 'V' 76-1259, 'E' 1261-2351 <RE2>
A:Cross-references: GB:X01740; MID:G182802; PIDN:AAA52484.1; PID:G182803
R:Frutect, M.A.; Blacher, R.; Burke, R.B.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo,
J.; Randolph, A.; Ordea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; Nd
DNA 4, 333-349, 1985
A:Title: Characterization of the polypeptide composition of human factor VIII:C and the
A:Reference number: A23584; PMID:86081164; PMID:3935400
A:Accession: A23584
A:Molecule type: mRNA
A:Residues: 1-2351 <TRU>
A:Cross-references: GB:M14113; NID:G182817; PIDN:AAA52485.1; PID:G182818
R:Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages
ity.
A:Reference number: A26174; PMID:86159740; PMID:3082357
A:Accession: A26174
A:Molecule type: protein
A:Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <EAT>
R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.
Biochemistry 31, 3315-3325, 1992
A:Title: Identification and functional importance of tyrosine sulfate residues within re
A:Reference number: A42346; PMID:92207952; PMID:1554716
A:Accession: A42346
A:Molecule type: protein
A:Residues: 20-36; 356-371; 392-408; 582-594; 1668-1669, 'X', 1671; 1672-1692; 1693-1708; 1709-17
A:Experimental source: recombinant material from Chinese hamster ovary cells
A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
R:Pay, P.J.; Smudzyn, T.M.
J. Biol. Chem. 264, 14005-14010, 1989
A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
A:Reference number: A43986; PMID:89340500; PMID:2503509
A:Accession: A43986
A:Molecule type: protein
A:Residues: 'X', 517-523; 1853-1860, 'X', 1862-1864, 'X', 1866 <FAY>
R:Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A:Title: Sulfation of Tyr(1690) of human blood coagulation factor VIII is essential for
A:Reference number: A56109; PMID:91093266; PMID:1898735
A:Contents: annotation; sulfation
R:Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A:Title: Characterization of the human factor VIII gene.
A:Reference number: A56196; PMID:85061547; PMID:6438525
A:Contents: annotation; introns
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o
A:Reference number: A56216; PMID:95338127; PMID:7613471
A:Contents: annotation; disulfide bonds
A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
R:Kjalkre, M.; Hedding, A.; Talbot, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
A:Reference number: S63527; PMID:96163459; PMID:8575434
A:Accession: S63527
A:Molecule type: protein
A:Residues: 733-752; 753-759 <KUA>
R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baekman, M.; Almstedt, A.; Gray, E.; Sandberg
Eur. J. Biochem. 232, 19-27, 1995
A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction

A:Reference number: S66445; PMID:96048024; PMID:7556150
A:Accession: S66445
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1668-1685 <LIN>
C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
C:Genetics:
A:Gene: GDB:F8C
A:Cross-references: GDB:119124; OMIM:306700
A:Map position: Xq28-Xq28
A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63;
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa prot
A:Pathway: blood coagulation
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; plas
F:1-19/DNA: signal sequence #status predicted <SIG>
F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
F:20-740/Product: coagulation factor VIIIa heavy chain #status experimental <ACH>
F:20-356/DNA: A1 <DA1>
F:23-348/DNA: ferroxidase repeat homology <FO1>
F:392-759/DNA: A2 <DA2>
F:402-730/DNA: ferroxidase repeat homology <FO2>
F:760-1667/DNA: B <DB0>
F:1668-2351/Product: coagulation factor VIIIa light chain #status experimental <ACL>
F:1709-2038/DNA: A3 <DA3>
F:1716-2038/DNA: ferroxidase repeat homology <FO3>
F:2039-2191/DNA: C1 <DC1>
F:2039-2188/DNA: discoidin I amino-terminal homology <DNI>
F:2192-2351/DNA: C2 <DC2>
F:2192-2345/DNA: discoidin I amino-terminal homology <DN2>
F:60-288, 601, 776, 803, 847, 919, 962, 982, 1020, 1024, 1074, 1085, 1204, 1274, 1278, 1301, 1319, 1403, 14
F:172-158, 267-348, 547-573, 649-730, 1851-1877, 1918-1922, 2040-2188/disulfide bonds: #status
F:365-336/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
F:365-737/Cleavage site: 1683, 1699/Binding site: sulfate (Tyr) (covalent) #status experimental
F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:414, 426/Binding site: sulfate (Tyr) (covalent) #status predicted
F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F:2193-2345/disulfide bonds: #status predicted

Query Match 100.0% Score 484; DB 1; Length 2351;
Best Local Similarity 100.0% Pred. No. 4.7e-44;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VNNPKWQLQVDFQKTMKVTGVTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 60
Db 2242 VNNPKWQLQVDFQKTMKVTGVTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKV 2301
Qy 61 FQGNQDSFPVNSLDPLLTLYLRHPQSW 91
Db 2302 FQGNQDSFPVNSLDPLLTLYLRHPQSW 2332

RESULT 3
A47004
coagulation factor VIII precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A47004
R:Elder, B.; Lakich, D.; Gitschier, J.
Genomics 16, 374-379, 1993
A:Title: Sequence of the murine factor VIII cDNA.
A:Reference number: A47004; PMID:93300511; PMID:8314577
A:Accession: A47004
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2319 <ELD>
A:Cross-references: UNIPROT:O06194; GB:I05573; NID:G192456; PIDN:AAA37385.1; PID:G192457
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F:1-19/DNA: signal sequence #status predicted <SIG>

F:23-349/Domain: ferroxidase repeat homology <FOI>
F:402-730/Domain: ferroxidase repeat homology <FO2>
F:1686-2006/Domain: ferroxidase repeat homology <FO3>
F:2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F:2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 83.9%; Score 406; DB 2; Length 2319;
Best Local Similarity 82.4%; Pred. No. 1.5e-35;
Matches 75; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 VNNKRWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 2210 VNDPKQWLQVLDLQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 2269

Qy 61 FQGNQDSFTPVVNSLDPPLLRHLPQSW 91
Db 2270 FQGNQDSFTPVVNSLDPPLLRHLPQSW 2300

RESULT 4
T42763
coagulation factor VIII precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42763
R:Lollar, P.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z22269
A:Accession: T42763
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2133 <LOL>
A:Cross-references: UNIPROT:P12263; EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AA8067
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 81.0%; Score 392; DB 2; Length 2133;
Best Local Similarity 79.1%; Pred. No. 4.5e-34;
Matches 72; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 1 VNNKRWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 60
Db 2024 VSSAEWLQVLDLQKTMKVTGTTQGVKSLTSMYVKEFLVSSQDGHQWTLFFQNGKVKV 2083

Qy 61 FQGNQDSFTPVVNSLDPPLLRHLPQSW 91
Db 2084 FQGNQDSFTPVVNSLDPPLLRHLPQSW 2114

RESULT 5
JC4915
ags protein precursor - rat
N:Alternate names: O-acetyl-Gd3 ganglioside
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: JC4915
R:Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.
Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A:Reference number: JC4915; MUID:96374422; PMID:8780713
A:Accession: JC4915
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-427 <OGU>
A:Cross-references: UNIPROT:P70490; DDBJ:D84068; NID:g1620006; PIDN:BAAL2210.1; PID:g162
A:Experimental source: CST cell
C:Comment: This protein is required for the O-acetylation of disialoganglioside sialic a
C:Genetics:

A:Gene: ags
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolog
F:1-21/Domain: signal sequence #status predicted <SIG>
F:28-60/Domain: EGF homology <EG1>
F:68-107/Domain: EGF homology <EG2>
F:110-267/Domain: discoidin I amino-terminal homology <DN1>
F:271-427/Domain: discoidin I amino-terminal homology <DN2>

Query Match 43.0%; Score 208; DB 2; Length 427;
Best Local Similarity 43.3%; Pred. No. 7.6e-15;
Matches 39; Conservative 15; Mismatches 36; Indels 0; Gaps 0;

Qy 2 NNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 61
Db 325 NSAKEWLQVLDLQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 384

Qy 62 QGNQDSFTPVVNSLDPPLLRHLPQSW 91
Db 385 QGNLNNSHKKNIFEKPFMARYVRLPLSW 414

RESULT 6
A36479
milk fat globule membrane protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A36479
R:Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzuki, D.; Srinivasan, U.; Parry, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the exist
A:Reference number: A36479; MUID:91046008; PMID:2122462
A:Accession: A36479
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-463 <STU>
A:Cross-references: UNIPROT:P21956; GB:M39337; NID:g19142; PIDN:AAA39534.1; PID:g199143
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolog
C:Keywords: membrane protein
F:28-60/Domain: EGF homology <EG1>
F:68-107/Domain: EGF homology <EG2>
F:147-303/Domain: discoidin I amino-terminal homology <DN1>
F:307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 42.8%; Score 207; DB 1; Length 463;
Best Local Similarity 42.2%; Pred. No. 1.1e-14;
Matches 38; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

Qy 2 NNPKEWLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 61
Db 361 NSAKEWLQVLDLQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKV 420

Qy 62 QGNQDSFTPVVNSLDPPLLRHLPQSW 91
Db 421 QGNLNNSHKKNIFEKPFMARYVRLPLSW 450

RESULT 7
T42764
coagulation factor V - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42764
R:Yang, T.L.; Cui, J.; Rehmtulla, A.; Yang, A.; Mousalli, M.; Kaufman, R.J.; Ginsburg,
Blood 91, 4593-4599, 1998
A:Title: The structure and function of murine factor V and its inactivation by protein C.
A:Reference number: Z22270; MUID:98282202; PMID:9616155
A:Accession: T42764
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2183 <VAN>
A:Cross-references: UNIPROT:O88783; EMBL:U52925; NID:g3219690; PID:g3219691; PIDN:AAC995;
C:Function: blood coagulation

C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma; f
F:350-682/Domain: ferroxidase repeat homology <FOX1>
F:1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 42.6%; Score 206; DB 2; Length 2183;
Best Local Similarity 42.4%; Pred. No. 9e-14;
Matches 39; Conservative 19; Mismatches 32; Indels 2; Gaps 1;

QY 2 NNPKWQLQVDFQKMTKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV--K 59
DB 2076 NNNKQWLQVLLKIKVTAIVTQGCSSLSEMYKSVIYQSGQVAMKPYKQKSSMVDK 2135

QY 60 VFQGNQDSFTPVNSLDPPLLTLYRLRHPSW 91
DB 2136 IFEGNNTKGHVKNFNPFIISRFIRIPKTW 2167

RESULT 8
KF05
N:Alternate names: coagulation factor V precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1989 #sequence, revision 02-Jun-1995 #text_change 08-Dec-2000
C:Accession: A56172; A42344; A28028; A27498; A25897
R:Crripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A:Title: Structure of the gene for human coagulation factor V.
A:Reference number: A42344; MUID:92232668; PMID:1567832
A:Accession: A56172
A:Molecule type: DNA
A:Residues: 1-2224 <CR1>
A:Cross-references: GB:J05368
A:Accession: A42344
A:Molecule type: DNA
A:Residues: 48-58; 79-89; 120-130; 191-201; 239-249; 313-323; 368-378; 428-437; 461-471; 533-542;
2070; 2111-2120; 2172-2181 <CR2>
R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A:Title: Complete cDNA and derived amino acid sequence of human factor V.
A:Reference number: A28028; MUID:87260886; PMID:3110773
A:Accession: A28028
A:Molecule type: mRNA
A:Residues: 1-857; R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <JEN>
A:Cross-references: GB:M16967
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum
A:Reference number: A27498; MUID:88107560; PMID:2827731
A:Accession: A27498
A:Molecule type: mRNA
A:Residues: 1-1284, 'I', 1286-1600 <KAN>
A:Cross-references: GB:M17785
A:Note: parts of this sequence were determined by protein sequencing
R:Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog
A:Reference number: A25897; MUID:86313665; PMID:3092220
A:Accession: A25897
A:Molecule type: mRNA
A:Residues: 1188-1215; 1315-2224 <KA2>
A:Cross-references: GB:M14335
A:Note: parts of this sequence were determined by protein sequencing
R:Keller, F.G.; Ortel, T.L.; Quinn-Alten, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A:Title: Thrombin-catalyzed activation of recombinant human factor V.
A:Reference number: A56139; MUID:95210278; PMID:7696276
A:Ccontents: annotation; thrombin cleavage sites
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C:Genetics:
A:Gene: GDB:P5
A:Cross-references: GDB:119896; OMIM:227400

A:Map position: 1q23-1q23
A:Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1; 655
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prote
A:Pathway: blood coagulation
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma; f
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-224/Domain: signal sequence #status predicted <SIG>
F:29-737/Domain: coagulation factor V #status predicted <MAT>
F:29-345/Domain: A1 <DN1>
F:33-323/Domain: ferroxidase repeat homology <FO1>
F:346-691/Domain: A2 <DA2>
F:351-684/Domain: ferroxidase repeat homology <FO2>
F:692-1573/Domain: B <DOB>
F:1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F:1574-2224/Domain: coagulation factor Va light chain #status experimental <VAL>
F:1574-1905/Domain: A3 <DA3>
F:1581-1905/Domain: ferroxidase repeat homology <FO3>
F:1667-1765/Region: phospholipid binding #status predicted
F:1906-2064/Domain: C1 <DC1>
F:1906-2064/Domain: discoidin I amino-terminal homology <DN1>
F:2065-2224/Domain: C2 <DC2>
F:2065-2224/Domain: discoidin I amino-terminal homology <DN2>
F:51-55; 239-297; 460-468; 554-741; 752-760; 776-782; 821-938; 977-1074; 1083-1103; 1106-1479; 1495
F:167-193; 248-329; 500-526; 603-684; 1725-1751; 1907-2061; 2066-2221/bisulfide bonds: #status
F:334-335/cleavage site: Arg-Asn (protein C) #status predicted
F:363-693/cleavage site: Arg-Asn (protein C) #status predicted
F:372-377/cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F:378-379/cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F:382-383/cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F:534-535/cleavage site: Carboxylate (Asn) (covalent) #status absent
F:737-738/cleavage site: Arg-Gly (protein C) #status predicted
F:737-738/cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F:1046-1047/cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experimental
F:1573-1574/cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 41.1%; Score 199; DB 1; Length 2224;
Best Local Similarity 38.0%; Pred. No. 5.3e-13;
Matches 35; Conservative 24; Mismatches 31; Indels 2; Gaps 1;

QY 2 NNPKWQLQVDFQKMTKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKV--K 59
DB 2117 NNNKQWLQVLLKIKVTAIVTQGCSSLSEMYKSVIYQSGQVAMKPYKQKSSMVDK 2176

QY 60 VFQGNQDSFTPVNSLDPPLLTLYRLRHPSW 91
DB 2177 IFEGNNTKGHVKNFNPFIISRFIRIPKTW 2208

RESULT 9
KF05
coagulation factor V precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence, revision 28-Apr-1995 #text_change 09-Jul-2004
C:Accession: A42580; A36497
R:Guineto, E.R.; Esmon, C.T.; Mann, K.G.; MacGillivray, R.T.
J. Biol. Chem. 267, 2971-2978, 1992
A:Title: The complete cDNA sequence of bovine coagulation factor V.
A:Reference number: A42580; MUID:92147638; PMID:1737753
A:Accession: A42580
A:Molecule type: mRNA
A:Residues: 1-2211 <GUI>
A:Cross-references: UNIPROT:Q28107; GB:M81440; NID:g163037; PIDN:AAA30512.1; PID:g163038
A:Note: sequence extracted from NCBI backbone (NCBI:80774, NCBIP:80776)
R:Kalafatis, M.; Jenny, R.J.; Mann, K.G.
J. Biol. Chem. 265, 21580-21589, 1990
A:Title: Identification and characterization of a phospholipid-binding site of bovine fac
A:Reference number: A36497; MUID:91072354; PMID:2254316
A:Accession: A36497
A:Molecule type: protein
A:Residues: 1566-1570, 'X', 1572-1581, 'X', 1583-1584; 1673-1676, 'X', 1678-1679, 'X', 1681, 'X', 16
R:Xue, J.; Kalafatis, M.; Silvera, J.R.; Kung, C.; Mann, K.G.
Biochemistry 33, 13109-13116, 1994
A:Title: Determination of the disulfide bridges in factor Va heavy chain.

A;Reference number: A55979; MUID:95034740; PMID:7947716
A;Contents: annotation
A;Note: 566-Cys and 617-Cys were shown to have free sulfhydryls
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prothrombinase
C;Pathway: blood coagulation
C;Superfamily: coagulation factor V; disocoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-221/Product: coagulation factor V #status predicted <MAT>
F;29-741/Product: coagulation factor Va heavy chain #status predicted <VAH>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-695/Domain: A2 <DA2>
F;351-688/Domain: ferroxidase repeat homology <FO2>
F;696-1564/Domain: B <DOB>
F;1175-1437/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1565-2211/Product: coagulation factor Va light chain #status predicted <VAL>
F;1565-1892/Domain: A3 <DA3>
F;1572-1892/Domain: ferroxidase repeat homology <FO3>
F;1654-1752/Region: phospholipid binding #status predicted
F;1893-2051/Domain: C1 <DC1>
F;1893-2048/Domain: disocoidin I amino-terminal homology <DN1>
F;2052-2211/Domain: C2 <DC2>
F;2052-2208/Domain: disocoidin I amino-terminal homology <DN2>
F;167-193,248-329,499-525/Disulfide bonds: #status experimental
F;235,239,297,382,460,563,587,745,756,774,780,902,952,964,1044,1053,1062,1071,1078,1094,
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363,697,1537/Binding site: sulfate (Tyr) (covalent) #status predicted
F;337-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;533-534/Cleavage site: Arg-Gly (protein C) #status predicted
F;607-688,1712-1738,1894-2048,2053-2208/bisulfide bonds: #status predicted
F;741-742/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1034-1035/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1564-1565/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 40.9%; Score 198; DB 1; Length 2211;
Best Local Similarity 39.1%; Pred. No. 6.8e-13;
Matches 36; Conservative 20; Mismatches 34; Indels 2; Gaps 1;

Qy 2 NNPKELQVDFQKTKMKTGVTQTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKV--K 59
Db 2104 NNNQWLQIDLLKTKTKTAIVTQGSLSSEMYKSYTHYSQDQTKPKYREKSSMWDC 2163
Qy 60 VFQGNQDSFTPVVNSLDPPLLTLYRLRHPSQW 91
Db 2164 IFEGNNVRGHVKNFFNPPIISRIPIRIPKTW 2195

RESULT 10
PP47 protein - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11743
R;Enslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toepfe
Biol. Reprod. 58, 1057-1064, 1998
A;Title: Molecular cloning and characterization of P47, a novel boar sperm-associated zo
A;Reference number: Z17325; MUID:98206817; PMID:9546740
A;Accession: T11743
A;Molecule type: mRNA
A;Status: preliminary
A;Residues: 1-409 <ENS>
A;Cross-references: UNIPROT:P79385; EMBL:Y11683; NID:G2652927; PIDN:CAAT72379.1; PID:G265
A;Experimental source: testis
C;Function:
A;Description: may be involved in membrane remodeling and/or function as a zona pellucid
C;Superfamily: milk fat globule protein; disocoidin I amino-terminal homology; EGF homolo
F;6-40/Domain: EGF homology <EGF>

Query Match 36.0%; Score 174; DB 2; Length 409;
Best Local Similarity 37.0%; Pred. No. 3.7e-11;

Matches 34; Conservative 17; Mismatches 39; Indels 2; Gaps 1;
Qy 2 NNPKELQVDFQKTKMKTGVTQTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKV--K 59
Db 305 NSASEWLQIDLGSQRRVTGIIITQGDGFHGIQYVAAYKVAYSDDGVSTEVYRDQGALEBK 364
Qy 60 VFQGNQDSFTPVVNSLDPPLLTLYRLRHPSQW 91
Db 365 IFPGNLDNNSHKMMFETPFLTRFVRILPVAW 396

RESULT 11
JC5256
adipocyte transcription factor, AEBP1 - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: JC5256
R;Ohno, I.; Hashimoto, J.; Shimizu, K.; Takaoka, K.; Ochi, T.; Matsubara, K.; Okubo, K.
Biochem. Biophys. Res. Commun. 228, 411-414, 1996
A;Title: A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expressi
A;Reference number: JC5256; MUID:97079196; PMID:8920928
A;Accession: JC5256
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-845 <OHN>
A;Cross-references: UNIPROT:Q14113; DBJ:D86479; NID:G1468942; PIDN:BAAL3094.1; PID:G1461

Query Match 36.0%; Score 174; DB 2; Length 845;
Best Local Similarity 34.9%; Pred. No. 8.8e-11;
Matches 30; Conservative 20; Mismatches 36; Indels 0; Gaps 0;

Qy 6 EWLQVDFQKTKMKTGVTQTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVQFQNGQ 65
Db 129 QWIEVDTRRTTRFTGVTITQGRDSSIHDDFVTTFFVGFSDNSQTVMYTNNGVEEMTFHGNV 188
Qy 66 DSFTPVVNSLDPPLLTLYRLRHPSQW 91
Db 189 DKDTPVLSELPEPVVAFIRIYPLTW 214

RESULT 12
SS1739
transcription repressor AEBP1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S60227; S51739
R;He, G.P.; Muise, A.; Li, A.W.; Ro, H.S.
Nature 378, 92-96, 1995
A;Title: A eukaryotic transcriptional repressor with carboxypeptidase activity.
A;Reference number: S60227; MUID:96061010; PMID:7477299
A;Accession: S60227
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-719 <HE2>
A;Cross-references: UNIPROT:Q61281; EMBL:X80478; NID:G507131; PIDN:CAAS56648.1; PID:G60713

Query Match 35.7%; Score 173; DB 2; Length 719;
Best Local Similarity 34.9%; Pred. No. 9.4e-11;
Matches 30; Conservative 20; Mismatches 36; Indels 0; Gaps 0;

Qy 6 EWLQVDFQKTKMKTGVTQTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVQFQNGQ 65
Db 24 QWIEVDTRRTTRFTGVTITQGRDSSIHDDFVTTFFVGFSDNSQTVMYTNNGVEEMTFHGNV 83
Qy 66 DSFTPVVNSLDPPLLTLYRLRHPSQW 91
Db 84 DKDTPVLSELPEPVVAFIRIYPLTW 109

RESULT 13
A47285
milk fat globule protein - human (fragment)
C;Species: Homo sapiens (man)

C; Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 C; Accession: A47285
 R; Jarocca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
 A; Title: A M-r 46,000 human milk fat globule protein that is highly expressed in human h
 A; Reference number: A47285; MUID:91371351; PMID:1909932
 A; Accession: A47285
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-218 <LAR>
 A; Cross-references: UNIPROT:Q08431; GB:S56151; NID:9235396; PIDN:AB19771.1; PID:9235397
 A; Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 F; 1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
 F; 60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match 33.9%; Score 164; DB 2; Length 218;
 Best Local Similarity 35.4%; Pred. No. 2.1e-10;
 Matches 31; Conservative 20; Mismatches 35; Indels 2; Gaps 1;

QY 6 EWLQVDQKMTKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFF--QNGKVKVFGQ 63
 Db 118 QMLQVDLGSSEKVTGIIQTGARNFGSVQFVASYKVAYNSANWTEYODPTRTGSSKIFPG 177

QY 64 NQDSFTFVNLSLDPPLLTLYRLHPSQW 91
 Db 178 NWDNHSKKNLFTPTILARYVRLPVAW 205

RESULT 14
 S65138
 Glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
 N; Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 h
 C; Species: Bos primigenius taurus (cattle)
 C; Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
 C; Accession: S65138; MUID:96125736; PMID:8541316
 R; Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuura, T.
 Biochim. Biophys. Acta 1245, 385-391, 1995
 A; Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal an
 A; Reference number: S65138; MUID:96125736; PMID:8541316
 A; Status: preliminary; not compared with conceptual translation
 A; Molecule type: mRNA
 A; Residues: 1-401 <AOK>
 R; Mather, I.H.; Baughart, L.R.; Lane, W.S.
 Biochem. Mol. Biol. Int. 29, 545-554, 1993
 A; Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
 Ii-like sequences.
 A; Reference number: A48394; MUID:93250576; PMID:8485470
 A; Accession: G48394
 A; Status: preliminary
 A; Molecule type: protein
 A; Residues: 207-220 <MAT>
 A; Experimental source: milk
 A; Note: sequence extracted from NCBI backbone (NCBIP:131457)
 C; Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 F; 1-18/Domain: signal sequence #status predicted <SIG>
 F; 19-427/Product: PAS-6/7 protein #status experimental <MAT>
 F; 24-58/Domain: EGF homology <EG1>
 F; 66-105/Domain: EGF homology <EG2>
 F; 108-265/Domain: discoidin I amino-terminal homology <DN1>
 F; 269-427/Domain: discoidin I amino-terminal homology <DN2>
 F; 29-47,49-58,66-77,71-94,98-105/Disulfide bonds: #status predicted
 F; 2/3/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F; 34/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F; 59,227/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F; 109-265,252-256,270-427/Disulfide bonds: #status experimental

Query Match 32.9%; Score 159; DB 2; Length 401;
 Best Local Similarity 33.7%; Pred. No. 1.7e-09;
 Matches 31; Conservative 17; Mismatches 42; Indels 2; Gaps 1;

QY 2 NNPKEWLQVDQKMTKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNG--KVK 59
 Db 323 NSASEWLQIDLGSKRVTGIIQTGARDFGHIQVAAVYVAGDDGVVTWTEYKDPGASESK 382

QY 60 VFGQNDSTFVNLSLDPPLLTLYRLHPSQW 91
 Db 357 IFPGNMDNNSHKKNIFETPTFQARFVRIQVPAW 388

Search completed: April 20, 2005, 17:21:43

THIS PAGE BLANK (USPTO)